# The Ubiquitin Ligase \(CRL4^{Cdt2}\) Targets Thymine DNA Glycosylase for Destruction during DNA Replication and Repair

Johannes C. Walter Tamara Jeannine Slenn

The ubiquitin ligase $CRL4^{Cdt2}$ targets thymine DNA glycosylase for destruction during DNA replication and repair

## *Abstract*

The E3 ubiquitin ligase $CRL4^{Cdt2}$ targets proteins for destruction during DNA replication and following DNA damage (Havens and Walter, 2011). Its substrates contain "PIP degrons" that mediate substrate binding to the processivity factor PCNA at replication forks and damage sites. The resulting PCNA-PIP degron complex forms a docking site for $CRL4^{Cdt2}$, which ubiquitylates the substrate on chromatin. Several $CRL4^{Cdt2}$ substrates are known, including Cdt1, multiple CDK inhibitors, *Drosophila* E2f1, human Set8, *S. pombe* Spd1, and *C. elegans* Polη (Havens and Walter, 2011). An emerging theme is that $CRL4^{Cdt2}$ targets proteins whose presence in S phase is toxic.

Here, I used *Xenopus* egg extract to characterize a new $CRL4^{Cdt2}$ substrate, thymine DNA glycosylase (TDG). TDG is a base excision repair protein that targets G-U and G-T mispairs, which arise from cytosine and 5-methylcytosine deamination (Cortázar et al., 2007). Thus, TDG may function in epigenetic gene regulation via DNA demethylation, in addition to its canonical DNA repair function. A yet unknown E3 ubiquitin ligase triggers TDG destruction during S phase (Hardeland et al., 2007). Understanding TDG proteolysis in S phase is relevant to the regulation of DNA replication, DNA repair, and epigenetic control of gene expression.

I discovered that TDG contains a variant of the "PIP degron" consensus and that TDG is ubiquitylated and destroyed in a PCNA-, Cdt2-, and degron-specific manner during DNA repair and DNA replication in *Xenopus* egg extract. I further characterized what features of TDG contribute to its proteolysis. Interestingly, I could not identify any defects during DNA replication or during *Xenopus* embryonic development in response to a non-degradable form of TDG.

Additionally, I examined how interactions between $CRL4^{Cdt2}$ and multiple subunits of the PCNA homotrimer contribute to $CRL4^{Cdt2}$ function. In a popular model, PCNA functions as a

"tool belt" on DNA, binding three separate proteins through its individual subunits to facilitate rapid exchange of DNA replication and repair proteins as they are needed on DNA. To address this model, I generated a single chain polypeptide with three PCNA subunits connected through flexible linker sequences. I used this tool to determine how multiple PCNA subunits contribute to $CRL4^{Cdt2}$ function. I found that a single wildtype subunit is sufficient for modest destruction of the $CRL4^{Cdt2}$ substrate Cdt1, but complete Cdt1 destruction requires two separate wildtype subunits. Additionally, a single subunit was sufficient for leading strand elongation, challenging the "tool belt" model during DNA replication. I also discuss implications and future use of the single-chain PCNA.

# *Table of Contents*

# *List of Figures and Tables*

# *Acknowledgements*

Without the support of so many people, this entire endeavor would not have been possible. I have been fortunate to have so much personal and scientific support throughout my life and career that gave me the strength and aptitude to succeed.

There are countless people to thank for their contributions toward my time in graduate school:

First and foremost, I thank **Margie and Kurt Slenn** and the rest of **my immediate and extended family** for always encouraging me to succeed and for supporting my career decisions, even when they didn't understand them. I am grateful to them for shaping my inquisitive mind from a young age, for molding me into an assertive and resilient adult, for taking an interest in my work, and for offering career and life advice whenever I invite it.

I must thank my PhD advisor, **Johannes C. Walter**, for invaluable guidance to help me progress through my projects in the Walter lab. I thank Johannes for his high standards on data quality, his lengthy weekly discussions about our theories and ideas that guided this work to completion, and for allowing me to work on a number of challenging yet fascinating projects during my time in lab. Finally, I am forever indebted to Johannes for taking me in as a third year graduate student when my original lab was moving across the country and for prioritizing my timely completion of graduate school.

I am grateful to many members of the Walter lab, past and present, especially **Jieqiong Zhang (JQ)**, without whom I might have lost my sanity many times, for being a great friend, listener, and supporter when experiments failed as well as succeeded, and also for being one of the all-around nicest people I know. I am also forever grateful to **Christina Kam**, "the teenager," for chitchatting about life, science, and silly things. I am also thankful to her for helping with little things in lab so I could leave at night and for finding numerous pieces of information and data for me while I was writing from afar. I am appreciative to both JQ and Christina for group text messages, banter, and lab gossip.

Additionally, I thank **David Long** and **Nadia Shobnam**, my lunch buddies during my early years in the Walter lab, for their support and encouragement when I couldn't see the light at the end of the tunnel. I am appreciative of **Courtney Havens** for generating much of the foundation on which my project was based and for her scientific guidance, advice, and constructive criticism. I also want to thank the rest of the Walter lab, past and present—**Magda Budzowska**, **Ben Morris**, **Julien Duxin**, **James Dewar**, **Ravi Amunugama**, **Hasan Yardimci**, **Anna Loveland**,

**Yu Fu**, **Thomas Graham**, **Yuko Takai**, and visiting professor **Nick Rhind** for their scientific discussions, constructive criticism, goofy entertainment, and celebrations over the most minute successes.

Before joining the Walter lab, I had great experiences under the guidance of my original advisor, **W. Matthew Michael**. I thank Matt for respecting me as a scientist even early in my career, for allowing me to shape my own project, and for promoting a sense of work-life balance. I owe several previous Michael lab members for imparting the most important lesson of all during graduate school—to let it go when something goes wrong and focus on what to do next. I am obliged to **Chis Murphy** for guidance and friendship during the early stepping-stones of graduate school as well as advice on the final hurdle: the dissertation and defense. I am grateful to **Shan Yan** for being a great baymate, for encouraging me scientifically, and for sharing stories about his family and listening to mine. I thank **Angela Coveny-Holmes** for always putting things in perspective when life seemed upside down. And last but not least, I am forever beholden to **Ash Williams** for his scientific advice, attention to detail, friendly attitude in a sometimes callous scientific world, weekly Berryline trips, and most importantly for multiple critical readings and comments on this entire dissertation during the final stages of writing.

I would also like to thank the Kirschner lab for their hospitality and assistance while I tried to learn the complicated method of microinjections at the true eleventh hour, especially **Bob Freeman**, who taught me the new skills and explained the details of frog development, and **Allon Klein**, who was always around to answer questions as they arose. In addition, I want to thank the Loparo lab, especially Professor **Joe Loparo**, for being excellent neighbors and lunch buddies and for providing additional support.

I have been motivated and inspired by my dissertation advisory committee—Professors **Andrew Murray**, **Briana Burton**, and **Randy King**—who provided crucial advice regarding both my project and progression through graduate school. I thank them for thinking seriously about my project and offering useful and constructive guidance and criticism. I am grateful to them for keeping my life goals in mind as well as my scientific goals and for helping me progress through graduate school.

One of my most valuable experiences during graduate school was my involvement with Science in the News (SITN), a graduate student group focused on communicating science to the general public. I thank SITN for giving me a sense of direction during graduate school and for shaping my career goals. I am especially thankful to SITN co-directors and other board members during my involvement with the group for building and improving the organization and being supportive

of my own scientific endeavors as well, especially **Marshall Thomas**, **Jeff Teigler**, **Amy Gilson**, **Daisy Robinton**, **Kelsey Taylor**, **Ann Fiegen**, and **Morgan Thompson**.

Additionally, I'd like to acknowledge my MCB friends and classmates who helped me adjust to graduate school and provided support throughout, especially **Emily Gleason**, **Lori Huberman**, **Jennifer Jocz Tan**, **Kah Yong Tan**, **Derek Lau**, and **Ann Ran**.

I am appreciative of the many teachers and professors who sparked my initial interests in biochemistry and science in general during my middle school, high school, and college years, especially **Jim McCusker**, **Paul McGibney**, **Marcia Kilcoyne**, **Laura Runyen-Janecky**, **Jonathan Dattelbaum**, and **Bill Myers.** I am particularly thankful to my undergraduate research advisors **Valerie Kish** at the University of Richmond and **Dennis Kolson** at the University of Pennsylvania for shaping me as a scientist and encouraging me to pursue my teaching and research goals.

I could not have done any of this without my closest friends who have supported me during graduate school and tolerated me when my stress got the best of me. I thank **Anastasia Kharlamova Kass**, **Ashley Miller**, **Jenn Carman Daniel**, **Elizabeth Danka**, **Veronica McMahon**, **Katie Riedo**, and **Sarah Remmert** for allowing me to get away from lab without even physically leaving through constant email contact. I knew I could always count on a near immediate response when I had good news to share or was having a bad day. I additionally thank Elizabeth Danka for her critical reading of my dissertation. Moreover, I thank **Jen Williams** for being a great friend and roommate during graduate school.

Last but not least, I am immeasurably thankful for the support of **Benjamin Lloyd**—for uncomplainingly absorbing the brunt of my frustrations toward science, graduate school, and life, for reminding me of my strengths and assets, for putting me first despite his incredibly busy schedule, for keeping me laughing when times got tough, and for patiently waiting for me to finish graduate school to start our life together.

# Chapter 1

# Background

## *1.1 DNA replication and repair*

All eukaryotic cells undergo the same basic cell cycle progression, and many cellular processes are highly conserved across all species of the three domains of life. After cell division, cells enter their first gap phase, G1, during which they grow, metabolize energy, and prepare for the subsequent phase. After G1, cells enter a synthesis phase, S phase, in which they generate cellular organelles and copy their entire genome exactly once. After S phase, cells enter their second gap phase, G2, and prepare for the act of cell division. Finally, in mitosis (M phase), the nuclear envelope dissolves to allow for equal DNA segregation between two daughter cells. As new nuclear envelopes form, the cellular membrane invaginates to create two new, distinct cells.

Many important events critical for progression through the cell cycle occur during S phase, including DNA replication, which is the focal point of the work described here. During DNA replication, each cell must make one complete copy of its genome and epigenome (inheritance of some epigenetic modifications is discussed in Chapter 1.3). Replication of a region of the genome more than once (re-replication) or failure to complete DNA replication leads to genomic instability, a hallmark of cancer. To ensure that exactly one round of DNA replication occurs during each S phase, cells employ multiple safeguards. First, to guarantee that cells complete replication before segregating their chromosomes during mitosis, checkpoint signaling pathways detect single-stranded DNA (ssDNA) to prevent progression into M phase before replication is complete. Additionally, cells prepare DNA for replication during G1 phase but actually replicate their DNA during S phase of the cell cycle. After replication is complete, cells cannot prepare for another round of DNA replication until after segregating their chromosomes into separate daughter cells. Finally, cells repair DNA damage throughout the cell cycle to allow for accurate DNA replication and transcription.

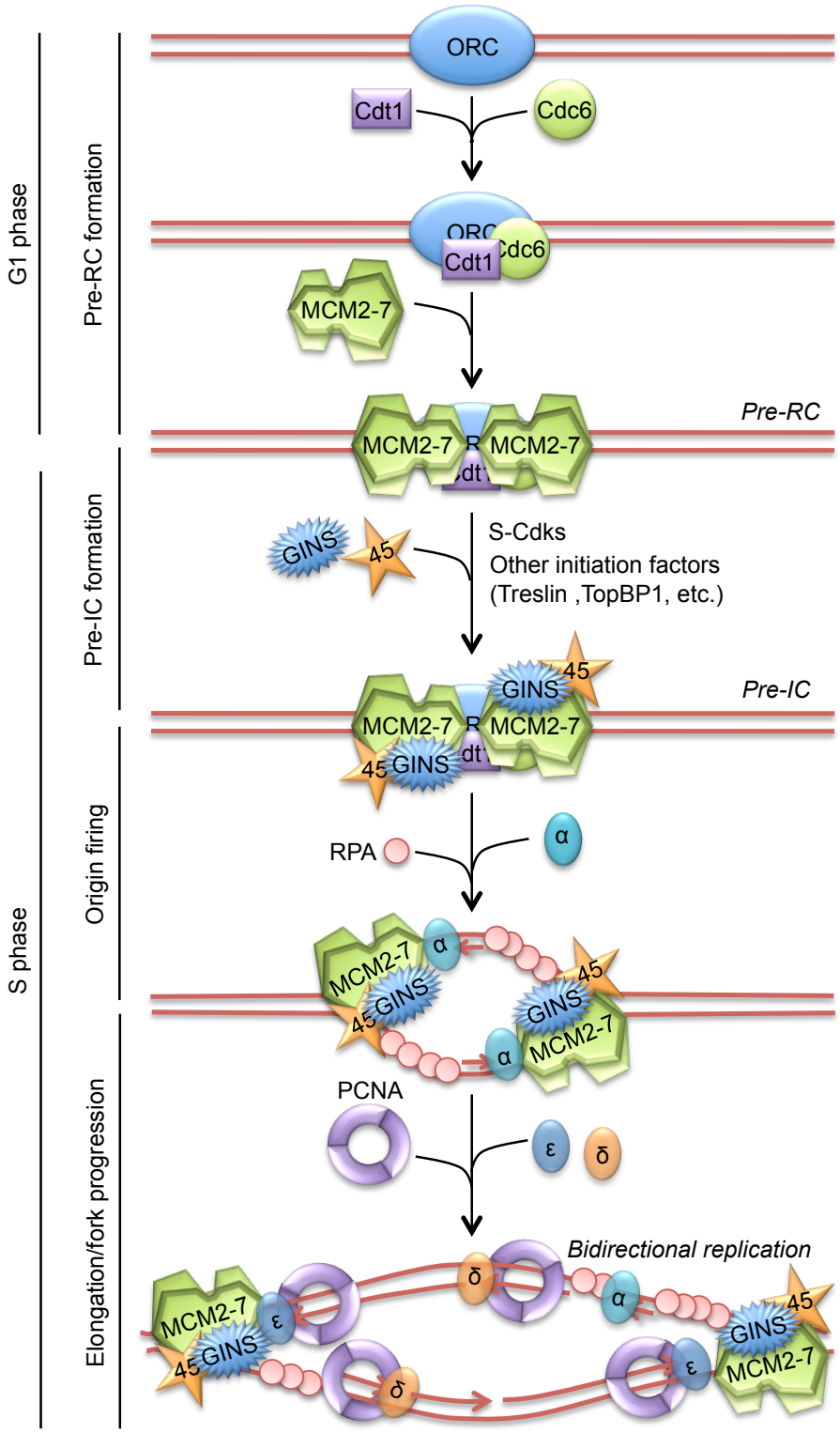
In all eukaryotes, DNA replication is remarkably similar, with conserved proteins and

mechanisms spanning from yeast to mammals (Sclafani and Holzen, 2007). A group of proteins called Cyclin-dependent kinases (Cdks), along with their partners, cyclins, control progression through the cell cycle. Although the complete repertoire of Cdks is typically present throughout the cell cycle, their corresponding partners, Cdk-specific cyclins, are synthesized and rapidly destroyed to control waves of Cdk activity (Murray, 1995; Murray and Kirschner, 1989). During G1, S phase-specific Cdk activity is low in nuclei so DNA replication cannot initiate, but cells prepare for DNA replication during this stage. Upon the transition to S phase, S phase-specific Cdk activity increases, and the chromatin decondenses to allow DNA replication (Heller et al., 2011; Masai et al., 2010). During these preparatory events in G1, termed "DNA licensing" (Blow and Laskey, 1988; Blow et al., 2011; Masai et al., 2010), the origin recognition complex (ORC) first assembles on DNA, either at sequence-defined origins in *S. cerevisiae* or at sequence-independent locations along the genome in most other eukaryotes (Cvetic and Walter, 2005). The ORC proteins then recruit the licensing factors Cdc6 (Cocker et al., 1996) and Cdt1 (Maiorano et al., 2000), which in turn recruit the MCM 2-7 helicase (Masai et al., 2010; Romanowski et al., 1996). Together, this complex is called the pre-replication complex (pre-RC). Once the MCM2-7 helicase has loaded, the DNA is considered to be licensed (Figure 1.1). Though the helicase is responsible for unwinding DNA during S phase to allow replication, it is inactive during G1 (Arias and Walter, 2007; 2004).

After entry into S phase, high levels of S phase-specific Cdks trigger the activation and relocalization of many proteins, including Cdc45, GINS, and other components of the "replisome progression complex" from the cytoplasm to the nucleus or from one complex to another within the nucleus (Arias and Walter, 2007; Masai et al., 2010) (Figure 1.1). Cdc45 and GINS associate with MCM2-7 to form the functional eukaryotic replicative helicase, referred to as the CMG (Cdc45, MCM2-7, and GINS) complex (Ilves et al., 2010; Makarova et al., 2012; Watase et al., 2012). This pre-initiation complex (pre-IC) is capable of unwinding DNA (Ilves et al., 2010)

**Figure 1.1: The events of DNA replication.** DNA is first licensed in G1 to form pre-RCs. Upon the transition to S phase, Cdk activity and a multitude of other factors convert the pre-RC to a pre-IC, thereby converting the inactive MCM2-7 helicase into a the functional CMG (Cdc45-MCM-GINS) replicative helicase. The helicase separates the two DNA strands, exposing single-stranded DNA that is coated by RPA, allowing origin firing to occur. Polα/primase then generates short RNA/DNA primers, allowing PCNA loading by RFC and processive DNA synthesis by Polδ and Polε. Polδ teams up with Polα, PCNA, flap endonuclease machinery, and DNA ligase to complete discontinuous lagging strand synthesis. Polε couples with PCNA to perform continuous leading strand synthesis.

**Figure 1.1** (continued)



(Figure 1.1). During origin firing, the helicase unwinds a portion of DNA, allowing additional replication proteins to load, including the ssDNA binding protein replication protein A (RPA), which stabilizes the unwound DNA and prevents rehybridization (Wold, 1997). Concurrently, the polymerase (Pol) Polα/primase complex generates short primers of newly synthesized RNA and DNA (Wold, 1997). Next, the replication processivity factor proliferating cell nuclear antigen (PCNA), often called the sliding clamp, can be loaded onto the junction of RPA-coated ssDNA and Polα-generated double-stranded DNA (Figure 1.1) by replication factor C (RFC). PCNA contains three identical subunits that encircle the DNA and serve as a binding platform for other replication proteins (Moldovan et al., 2007), including the leading and lagging strand DNA polymerases Polε and Polδ, respectively (Burgers, 2008; Nick McElhinny et al., 2008). These polymerases synthesize new DNA during the course of S phase. Because of the antiparallel structure of DNA and the unidirectionality of polymerase activity, the leading strand is synthesized continuously, while the lagging strand is synthesized discontinuously in short Okazaki fragments (Figure 1.1). Hundreds of other DNA replication and repair proteins also use PCNA as a scaffold to access DNA (Moldovan et al., 2007).

Importantly, pre-IC and pre-RC formation are temporally distinct; pre-RCs form during G1, and they are only converted to pre-ICs during S phase. New pre-RCs cannot form once replication has initiated. Cells accomplish this temporal exclusivity through several redundant mechanisms that prevent MCM2-7 recruitment during S phase. Among these devices are the activity of geminin (McGarry and Kirschner, 1998; Tada et al., 2001), a Cdt1 inhibitor that is not present during G1 but is active during the rest of the cell cycle, and the export of Cdc6 from the nucleus once replication has initiated (Kim and Kipreos, 2008; Saha et al., 1998). Additionally, Cdt1 is destroyed via replication-dependent proteolysis due to the ubiquitin ligase $CRL4^{Cdt2}$ after origins fire and replication elongation begins. These redundant mechanisms are important to prevent re-replication. If origin licensing continued during S phase, pre-RCs could assemble on newly

replicated DNA, leading to additional pre-IC formation and origin firing. Several other regulatory mechanisms exist to prevent ORC formation, change the localization of pre-RC components, and to regulate MCMs.

Various types of DNA damage that can impede replication fork progression must be repaired during either S phase or other stages of the cell cycle to allow for accurate replication and mitosis. Some types of damage, such as thymine dimers caused by UV irradiation, can be bypassed during DNA replication (Prakash et al., 2005; Sale et al., 2012). Other types of damage, like interstrand crosslinks and double-strand breaks, must be repaired before replication can progress. Two main repair pathways handle the smaller lesions that affect single nucleotides and stall the replication fork: nucleotide excision repair (NER) and base excision repair (BER). Though many of these lesions can be bypassed, repair is necessary to minimize mistakes during replication and transcription. NER involves the detection of bulky lesions by machinery that scans the genome for damage, either globally or at sites of transcription. After lesion detection, the DNA is unwound, stabilized by RPA, and incised by endonucleases to generate a ssDNA gap that is then filled by DNA polymerases in a PCNA-dependent manner (Kamileri et al., 2012). Finally, the gap is sealed by a DNA ligase. Several of the factors involved in NER are DNA replication proteins themselves, such as PCNA, RPA, and the replicative polymerases. The BER pathway repairs many aberrant DNA bases and modifications that are not recognized by NER. For example, BER restores guanine from 8-oxoguanine, which forms due to metabolically generated reactive oxygen species. Additionally, BER is responsible for the repair of certain DNA mispairs and for removal of aberrant bases. For instance, the glycosylases thymine DNA glycosylase (TDG) and Mbd4 initiate regeneration of correct G-C base pairs from G-T or G-U mispairs (Krokan et al., 2002; Sjolund et al., 2012). Several other glycosylases, such as the uracil-N-glycosylases UNG1 and UNG2, are also capable of repairing G-U mispairs (Krokan et al., 2002).

BER involves several steps. First, a specific glycosylase, like TDG or Mbd4, detects its cognate aberrant DNA base, modification, or mispair. The glycosylase hydrolyzes the N-glycosidic bond of the aberrant base from the sugar-phosphate backbone, generating an abasic site (also called an apurinic site or AP site). Next, an AP endonuclease incises the DNA sugar-phosphate backbone to generate a nick that is detected by a DNA polymerase: in short-patch BER, Polβ fills in the single-nucleotide gap and removes the remaining sugar-phosphate fragment, whereas in long-patch BER, PCNA-associated Polβ, Polδ, or Polε perform a strand displacement reaction that is subsequently processed and ligated (Kim and Wilson, 2012). According to a popular model, the substrate channeling or "hand-off" model for BER, each enzyme-product complex serves to recruit the next player in the pathway to ensure nicked and broken DNA is not left unrepaired. For example, the glycosylase-bound abasic site recruits the AP endonuclease so the vulnerable abasic site is never exposed (Prasad et al., 2011; 2010; Wilson and Kunkel, 2000).

PCNA directs many processes during DNA replication and repair, from polymerase processivity, to Okazaki fragment processing, to protein turnover during S phase. As described above, the homotrimeric PCNA forms a clamp around DNA, thereby providing a scaffold for proteins to associate with DNA and to interact with each other. Each subunit of PCNA contains an interdomain connector loop (IDCL) that can bind proteins containing a PCNA interacting peptide motif (PIP box) (Moldovan et al., 2007). This interaction is further defined in Section 1.2. Some groups have proposed that PCNA and other sliding clamps function as a "tool belts" that can bind multiple proteins at once to allow these proteins rapid access to DNA upon a change in conditions (Freudenthal et al., 2011; Indiani et al., 2005; Sutton, 2010; Zhuang and Ai, 2010). In this scenario, PCNA could bind three proteins necessary for one process or bind repair proteins during replication in order to be prepared in the case of a problem. Other groups have proposed a "dynamic hand-off model," in which DNA-associated proteins recruit the next protein in line for

a given pathway (Prosperi, 2006), similar to the “hand-off” model for BER. In this scenario, one complex forms a scaffold for the formation of another, and only one PCNA-bound protein is necessary at a time.

Despite the plethora of coordinated recruitment events, redundant regulatory pathways, and DNA repair pathways, mistakes do occur during DNA replication, and DNA damage can persist during S phase, leading to double-strand breaks or replication fork failures. DNA lesions that stall replicative polymerases activate the Ataxia-telangiectasia mutated and Rad3-related (ATR)-dependent checkpoint pathway. The ATR checkpoint pathway delays the cell cycle, prevents late origin firing, and stabilizes replication forks to prevent collapse. ATR, through its effector checkpoint kinase 1 (Chk1), causes a cell cycle delay by inhibiting mitotic Cdk activity. The role of ATR in replication fork stabilization and inhibition of late origin firing is not well understood. Several laboratories have focused on determining how damaged DNA activates the ATR kinase and how ATR activates Chk1. Current models suggest that the primary damage detection mechanism requires DNA replication (Lupardus et al., 2002; Stokes, 2002). When replicative polymerases stall, their associated helicases continue unwinding DNA, exposing regions of ssDNA (Byun, 2005). Replication likely restarts with the generation of a primer downstream of the stalled polymerase (Lopes et al., 2006), and the new 5’ DNA or RNA end is required for checkpoint activation (MacDougall et al., 2007; Yan and Michael, 2009). The resulting DNA structure contains a ssDNA gap adjacent to the 5’ end of a newly synthesized primer. Together, RPA-coated ssDNA and the 5’ end of the primer coordinate the assembly of core checkpoint signaling components, leading to the activation of ATR and Chk1. The Ataxia-telangiectasia mutated (ATM) checkpoint signaling pathway similarly delays cell cycle progression in response to double-strand breaks (Andreassen, 2005).

Mutations in any of these critical pathways responsible for transactions on DNA—DNA replication, DNA repair, or checkpoint signaling—are linked to genome instability and cancer. Advances in understanding the molecular details of these events could therefore be valuable to the fields of cancer biology and pharmacology.

## *1.2 PCNA-dependent proteolysis through the CRL4$^{Cdt2}$ ubiquitin ligase*

During DNA replication and repair, PCNA acts as a processivity factor, director, and scaffold for proteins that bind it through their PIP boxes. One class of PIP box-containing proteins is a set of substrates for the Cullin 4-RING-based ubiquitin ligase (CRL4$^{Cdt2}$), composed of the proteins Cul4, Ddb1, and the putative substrate receptor, Cdt2 (Abbas and Dutta, 2011). Unlike other ubiquitin ligases, CRL4$^{Cdt2}$ functions specifically during DNA replication or following DNA damage. CRL4$^{Cdt2}$ negatively regulates the stability of many proteins whose presence in S phase is detrimental. The best-characterized substrate of CRL4$^{Cdt2}$ is Cdt1 (Jin et al., 2006), which recruits the replicative MCM2-7 helicase during origin licensing. After cells progress into S phase, Cdt1 activity could facilitate detrimental origin licensing on regions of the genome that have already been replicated (Arias and Walter, 2005). To prevent aberrant licensing, CRL4$^{Cdt2}$ marks Cdt1 for destruction upon entry into S phase. CRL4$^{Cdt2}$ also targets Cdt1 for destruction in response to DNA damage, though the reasons are unclear (Hu et al., 2004). One potential explanation is that damage-dependent Cdt1 destruction might be necessary to suppress re-replication after checkpoint activation in G2. Checkpoint signaling inhibits mitotic Cdk activity, and mitotic Cdk activity is important to block pre-RC formation in mitosis (Arias and Walter, 2004).

Fewer than ten CRL4$^{Cdt2}$ substrates are known, but all have a role in DNA replication or repair (Havens and Walter, 2011). Other substrates of CRL4$^{Cdt2}$ include *Drosophila* E2f1(Shibutani et al., 2008), human Set8 (Centore et al., 2010), *Xenopus* Xic1 (Kim et al., 2010), *S. pombe* Spd1

(Liu et al., 2003; Salguero et al., 2012), *C. elegans* POLH-1 (the DNA Polη homologue) (Kim et al., 2010) and p21 in multiple organisms (Kim et al., 2008). Destruction of each $CRL4^{Cdt2}$ substrate helps to ensure proper DNA replication, and most $CRL4^{Cdt2}$ substrates are toxic during S phase. Addressing why each substrate must be destroyed has helped us understand important molecular details about DNA replication and repair. For example, the $CRL4^{Cdt2}$ substrate Set8 is a histone methyltransferase that methylates histone H4 on lysine 20 (H4K20). Failure to destroy Set8 during replication has been associated with aberrant H4K20 methylation at replication origins and re-replication (Abbas et al., 2010; Beck et al., 2012; Tardat et al., 2010). The link between re-replication and Set8 function is still not clear, but a prevailing model proposes that H4K20 methylation during DNA replication leads to relicensing (Beck et al., 2012). Before the discovery that Set8 is a $CRL4^{Cdt2}$ substrate, the relationship between H4K20 methylation and cell cycle progression was unknown. Now, the connection between H4K20 methylation and origin firing is an area of active study. Therefore, I expect that identification of novel $CRL4^{Cdt2}$ substrates will similarly improve our knowledge of DNA replication and repair, as well as help us gain more insight into the activities of new substrates.

$CRL4^{Cdt2}$ substrates must bind PCNA through their PIP boxes (Figure 1.2) to function as degradation targets (Arias and Walter, 2006); however, $CRL4^{Cdt2}$ does not target all PIP box-containing proteins for destruction. To the contrary, the majority of PCNA interacting proteins are stable when bound to PCNA; PCNA facilitates their recruitment and function on DNA (Moldovan et al., 2007). All known $CRL4^{Cdt2}$ substrates contain a basic residue four amino acids downstream of the PIP box (the B+4 position relative to the PIP box) (Havens and Walter, 2009) that is not present in stable PCNA interacting proteins (Figure 1.2). These features (a PIP box and "B+4" residue) create a "PIP degron." The binding surface for $CRL4^{Cdt2}$ includes residues on both the substrate and on DNA-bound PCNA (Figure 1.2) (Havens and Walter, 2009). Our current model proposes that the PCNA-bound substrate recruits $CRL4^{Cdt2}$ through contacts

**Figure 1.2: CRL4$^{Cdt2}$ substrates contain a "PIP degron." A)** Most CRL4$^{Cdt2}$ substrates contain a conserved PIP box (consensus sequence shown at the top), threonine and aspartate at positions 5 and 6 of the PIP box (a "TD motif"), and a basic residue four amino acids downstream of the PIP box (a "B+4 residue"). Together, these features constitute a "PIP degron," whereΨ represents any hydrophobic residue (I/L/V/M), and θ represents any aromatic residue (Y/F/W). Residues shown in green are part of the conserved PIP box, and residues shown in blue additionally contribute to the "PIP degron." **B)** The PIP degron residues interact with the IDCL of PCNA (shown in gray) to generate a binding surface for Cdt2, the substrate specificity factor of CRL4$^{Cdt2}$. The p21 PIP degron is shown, and the residue colors are the same as in (A). PCNA residues D122 and E124, thought to be involved in ligase recruitment (Havens et al., 2012), are highlighted in orange. This figure was generated in PyMol by Benjamin Morris of the Walter lab using PDB accession number 1AXC (Gulbis et al., 1996).

**Figure 1.2** (continued)

A

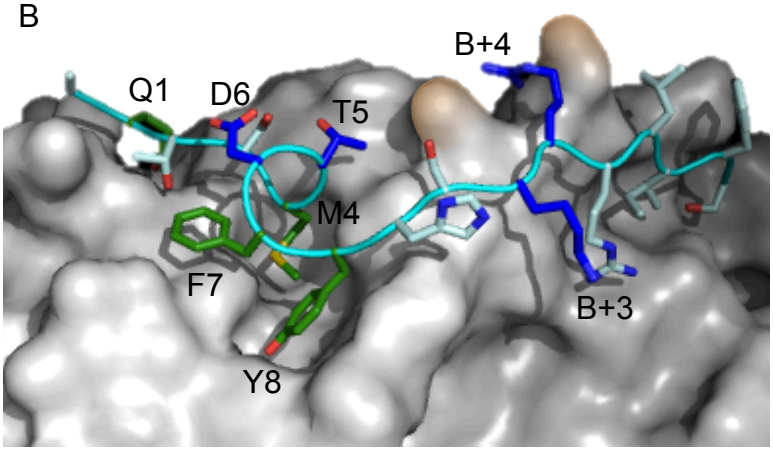
PIP box

12345678 +1 +2 +3 +4

Q/N XXΨXX θθXXXX

| | | |
|---|---|---|
| *Hs* Cdt1 | 3-QRRVTDFFARRR-14 | Confirmed CRL4$^{Cdt2}$ targets |
| *Hs* p21 | 144-QTSMTDFYHSKR-155 | |
| *Dm* E2F | 151-SNDITNYYKVKR-162 | |
| *Ce* POLH-1 | 572-PKSLESFFKKKK-583 | |
| *Hs* Set8 | 178-NRKLTDFYPVRR-189 | |
| *Xl* Xic1 | 170-TTPITDYFPKRK-181 | |

B



between Cdt2 and the PCNA-substrate degron, but that recruitment of the ubiquitin ligase does not require protein interactions on any PCNA subunits other than the substrate-bound subunit (Figure 1.3A). The ligase then polyubiquitylates the substrate, targeting it for proteasomal degradation. In an alternate model, Cdt2 recruitment requires its own interaction with another PCNA subunit in addition to its interaction with the substrate-bound PCNA subunit (Figure 1.3B).

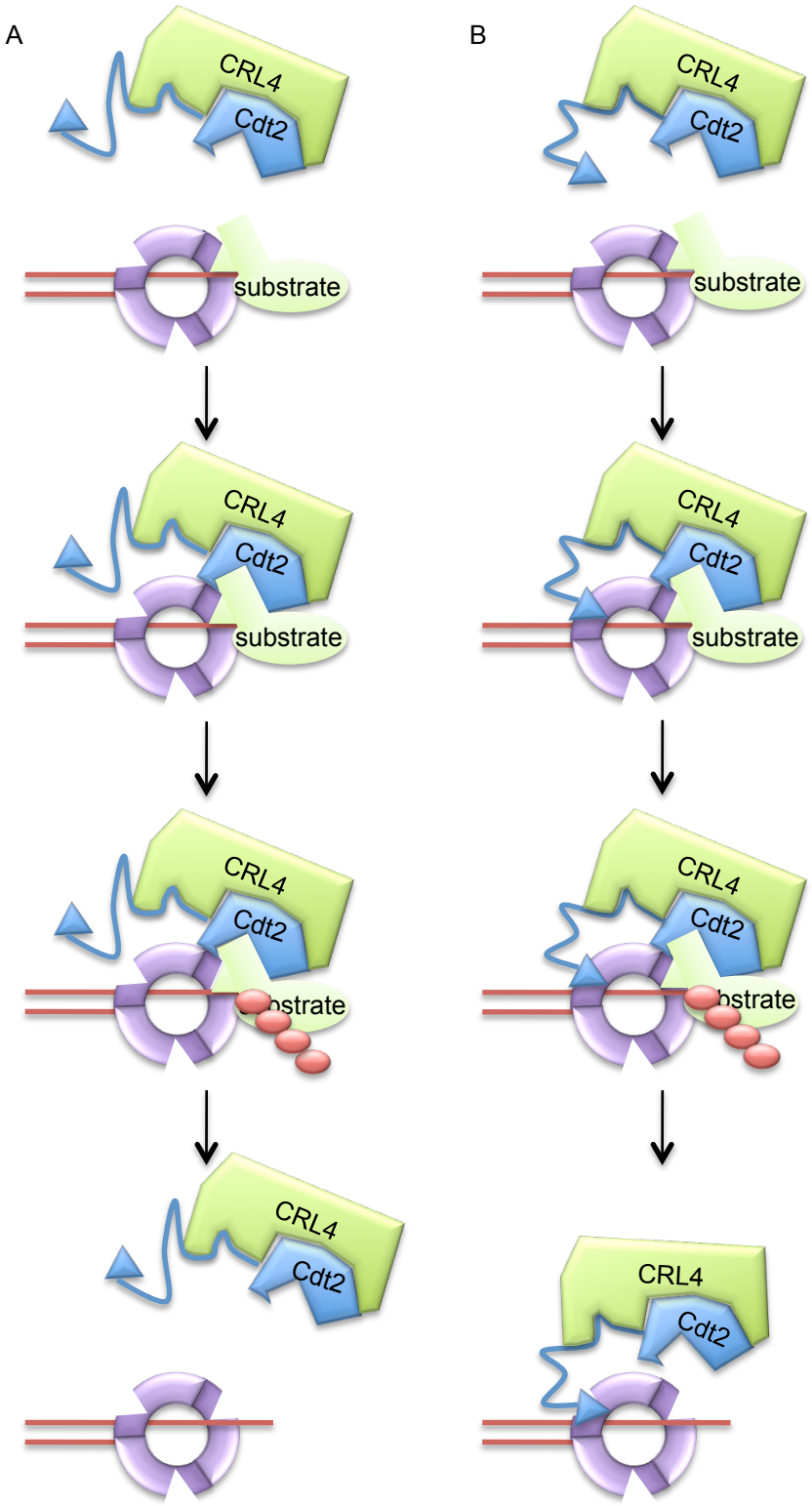
Though we have a general idea how $CRL4^{Cdt2}$ functions, many details are unclear. In the Walter lab, we are interested in identifying new substrates, expanding on the known sequence and structural requirements for substrates, and uncovering more details about the contributions of individual ligase components, the substrate, DNA, and PCNA to $CRL4^{Cdt2}$ function.

## *1.3 A role for base excision repair in DNA methylation*

DNA methylation is a critical epigenetic mark that controls gene expression. The DNA base cytosine can have a methyl group attached, generating 5-methylcytosine (5mC). Most often, animal cells methylate their DNA in CpG sequences (Chen and Riggs, 2011; Niehrs, 2009; Suzuki and Bird, 2008; Wu and Zhang, 2010). This alters chemical interactions between DNA and DNA binding proteins, leading to changes in chromatin compaction and transcription. Typically, methylation of gene promoters leads to gene silencing, though methylation of gene bodies has also been associated with gene activation (Chen and Riggs, 2011; Maunakea et al., 2010; Suzuki and Bird, 2008). Additionally, regions of heavily methylated DNA tend to exist as heterochromatin, a structure most often associated with repression of gene expression (Suzuki and Bird, 2008). A number of DNA methyltransferases (Dnmts) are specialized to either methylate unmodified DNA or to maintain its methylation status, particularly during DNA replication (Chen and Riggs, 2011; Wu and Zhang, 2010). Newly replicated DNA is hemimethylated: the parental strand retains its methylation status while the nascent strand is unmethylated. Maintenance Dnmts recognize this hemimethylated DNA and use it to direct their

**Figure 1.3: Models for CRL4$^{Cdt2}$ function. A)** The substrate binds to DNA-bound PCNA through its PIP box. The resulting complex then recruits the ubiquitin ligase through contacts between Cdt2 and the substrate-PCNA surface. CRL4$^{Cdt2}$ ubiquitylates the substrate, which is subsequently destroyed by the proteasome. This pathway is not processive: the ligase dissociates after one round of ubiquitylation. **B)** An alternate model for CRL4$^{Cdt2}$ function. This model is identical except that in this case, Cdt2 directly binds to a separate subunit on PCNA to facilitate its recruitment and substrate destruction. This could potentially lead to processive substrate destruction.

**Figure 1.3** (continued)



activity toward the nascent strand, thereby preserving the methylation status of individual genes through each round of DNA replication (Chen and Riggs, 2011; Jeltsch, 2006). Newly replicated DNA therefore inherits this sequence-independent epigenetic information through each cell division.

How cells initially decide to methylate or demethylate individual genes varies and is an area of active study. Although the key players in DNA methylation have been identified, proteins and pathways responsible for demethylation in animals are largely unknown. In plants, BER is responsible for DNA demethylation: several DNA glycosylases directly detect and hydrolyze methylated cytosine from the DNA backbone. After this initial base excision step, a cascade of BER steps occurs, resulting in the insertion of an unmethylated cytosine across from the guanine, thereby restoring the original unmethylated DNA sequence (He et al., 2011a; Wu and Zhang, 2010). In animals, however, no direct DNA glycosylase activity has been detected on 5mC, so the mechanism of DNA demethylation has remained a mystery. Some demethylation might occur passively; inhibiting the activity of maintenance methyltransferases toward hemimethylated DNA during replication leads to dilution of the epigenetic mark, so demethylation can occur over the course of multiple rounds of DNA replication (Niehrs, 2009). Passive DNA demethylation can be regulated at the level of maintenance Dnmts. However, many studies have observed active demethylation that does not require multiple rounds of replication. For example, in *X. laevis* oocytes, a reporter gene can be demethylated in the absence of DNA replication (Barreto et al., 2007; Simonsson and Gurdon, 2004). Notably, this demethylation does require DNA synthesis during DNA repair (Barreto et al., 2007; Wu and Zhang, 2010). Many examples of essential active demethylation have been described: mammalian zygotes rapidly demethylate the paternal genome before fusion of the maternal and paternal pronuclei (Iqbal et al., 2011; Mayer et al., 2000; Okada et al., 2010; Oswald et al., 2000; Wu and Zhang, 2010), cells of several developing organisms demethylate specific genes

during tissue differentiation (Ji et al., 2010; Kress et al., 2006; Wu and Zhang, 2010), and adult neurons demethylate certain genes to support memory and learning (Grayson and Guidotti, 2012; Guo et al., 2011; Leach et al., 2012; Niehrs and Schäfer, 2012).

Roles for both NER and BER have been proposed in discussions about DNA demethylation (Niehrs, 2009), but here I will focus on a putative role for BER in DNA demethylation. Several studies have identified a role for BER in transcriptional regulation through a role in DNA demethylation (Reviewed in Chen and Riggs, 2011; Nabel and Kohli, 2011; Niehrs, 2009; Sjolund et al., 2012; Wu and Zhang, 2010). Although the animal BER proteins TDG and MBD4 have no direct activity on 5mC, both act preferentially within large stretches of CpGs. This is not surprising since G-T mispairs often arise from spontaneous 5mC deamination or from the activity of other enzymes on 5mC; however, it is also possible that deamination and other modifications of 5mC occur in a controlled manner, leading to active demethylation upon their repair (Fritz and Papavasiliou, 2010; Niehrs, 2009; Wu and Zhang, 2010). A number of enzymes can act on 5mC to change this moiety to another DNA base. For example, the activation-induced cytosine deaminases AID and Apobec1 process 5mC directly to thymine, generating a G-T mispair (Morgan et al., 2004; Rai et al., 2008). Additionally, the more recently identified Ten eleven translocation (Tet) proteins iteratively oxidize 5mC to generate 5-hydroxymethylcytosine (5hmC) (Tahiliani et al., 2009), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC) (Ito et al., 2011). AID can also further process 5hmC to generate 5-hydroxymethyluracil (5hmU) (Guo et al., 2011). All of these products generated by 5mC-modifying enzymes could potentially serve as substrates for BER, allowing enzymes involved in the modification of 5mC to cooperate with BER to facilitate DNA demethylation. Notably, mutation of Tet proteins has been linked to both cancer and cellular differentiation defects (Cimmino et al., 2011; Williams et al., 2012).

## *1.4 Thymine DNA glycosylase*

TDG is a BER protein that binds to specific mispairs and aberrant bases in DNA such as G-T or G-U mispairs, uracil analogs like 5-fluorouracil (Kunz et al., 2009; Morgan et al., 2007), and 5-methylcytosine derivatives generated through Tet-mediated oxidation. After binding the mispair or aberrant base, TDG removes the incorrect base from the sugar-phosphate backbone through hydrolysis of the N-glycosidic bond (Cortázar et al., 2007). Most often, the aberrant base resides opposite a guanine in the complementary strand. Following base hydrolysis by TDG, AP endonuclease displaces TDG and cleaves the sugar-phosphate backbone so that DNA polymerase Polβ can insert cytosine across from guanine (Waters et al., 1999). TDG typically participates in short-patch BER, which does not involve a strand displacement reaction or PCNA (Kim and Wilson, 2012). Importantly, TDG binds virtually irreversibly to its abasic product; its removal depends on SUMOylation of the C-terminus, which decreases TDG's affinity for DNA (Baba et al., 2005; Hardeland et al., 2002; Smet-Nocca et al., 2011).

TDG also interacts with proteins involved in transcriptional regulation and has a role in epigenetic inheritance (Cortázar et al., 2007). In mice, TDG is essential, as $TDG^{-/-}$ embryos die during embryonic day 11.5, when cellular differentiation has just started (Cortázar et al., 2011; Cortellino et al., 2011; Saito et al., 2012). This phenotype appears to be a consequence of aberrant DNA methylation (Cortázar et al., 2011; Cortellino et al., 2011). In support of a direct role for TDG in DNA demethylation, a recent study showed that targeting TDG directly to specific genes results in demethylation of these genes (Gregory et al., 2012). Additionally, TDG physically interacts with the de novo methyltransferase Dnmt3a, though the purpose of this interaction is unclear (Li et al., 2006).

Two main models for TDG's role in demethylation have been proposed, both involving an initial modification of 5mC to another moiety followed by TDG activity. First, Tet proteins could oxidize

5mC to 5hmC and then to 5fC and 5caC. Though 5hmC is not a TDG substrate, both 5fC and 5caC are TDG substrates, and TDG-dependent BER of these modified bases would generate unmodified cytosine. In support of this first model, He and colleagues found that Tet2 catalyzes the formation of 5caC in human cells and that this modification accumulates in the absence of TDG (He et al., 2011b). Several other groups have expanded on this model and shown that Tet activity is important for demethylation of the paternal genome upon fertilization, that TDG can act on products of Tet enzymatic activity, and that Tet activity on methylated DNA also promotes DNA demethylation in adult tissue (Guo et al., 2011; Hashimoto et al., 2012; Inoue and Zhang, 2011; Inoue et al., 2011; Iqbal et al., 2011; Ito et al., 2011; Maiti and Drohat, 2011; Wossidlo et al., 2011; Zhang et al., 2012). This theory has been reviewed extensively (Cimmino et al., 2011; Nabel and Kohli, 2011; Williams et al., 2012). Alternatively, TDG could cooperate with AID or APOBEC, which generate G-T, G-U, or G-5hmU mispairs from G-5mC, G-C, and G-5hmC respectively. Again, TDG can act on each of these mispairs (Cortázar et al., 2007). In support of this second model, Cortellino and colleagues showed that TDG forms a complex with AID and Gadd45a, another protein implicated in DNA demethylation (Cortellino et al., 2011). Additionally, global erasure of DNA methylation during reprogramming of mouse primordial germ cells requires AID (Popp et al., 2010), and demethylation in zebrafish involves AID, Gadd45, and a glycosylase (Rai et al., 2008). It is possible that TDG cooperates with both Tet-mediated oxidation and AID-mediated deamination under different circumstances, but it is clear that TDG does have a fundamental role in DNA demethylation.

The domain organization and structure of TDG relevant to BER has been well characterized (Figure 1.4). The central catalytic core of TDG is conserved among DNA glycosylases and possesses a DNA-binding domain (Cortázar et al., 2007). Although the detailed atomic structure of the N-terminal regulatory domain has not been determined, it clearly has a role in DNA binding and undergoes structural rearrangement upon SUMOylation of the C-terminus

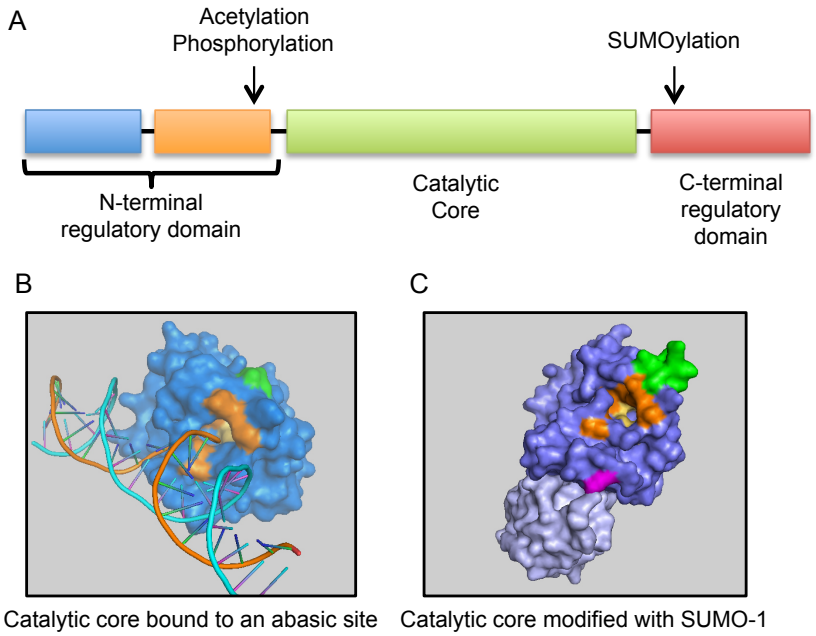



**Figure 1.4: The structure of TDG. A)** The domain structure of TDG. The N-terminal regulatory domain is unstructured, especially at the extreme N-terminus, shown in blue. TDG is both acetylated and phosphorylated in the N-terminal regulatory domain close to the catalytic core. The catalytic core shares sequence homology with other uracil DNA glycosylases. The C-terminal regulatory domain contains a SUMOylation site. **B)** TDG (catalytic core) is shown tightly bound to its catalytic product, the abasic site. This structure was generated in PyMol using PDB accession number 2RBA (Maiti et al., 2008). **C)** TDG (catalytic core) is shown upon modification with SUMO-1 (shown in light purple). This structure was generated in PyMol using PDB accession number 1WYW (Baba et al., 2005). For both (B) and (C), residues highlighted in orange are part of the catalytic site, and the residue highlighted in yellow is the critical asparagine in the active site. The region highlighted in green is where the N-terminus connects to the catalytic core. The N-terminus was not crystalized.

(Baba et al., 2005). The N-terminal region also contains several sites for modification, including acetylation and phosphorylation (Mohan et al., 2010). The C-terminus also serves a regulatory role and contains the SUMOylation site (Cortázar et al., 2007). A number of post-translational modifications—acetylation, phosphorylation, and SUMOylation—regulate the function of TDG in DNA repair, transcriptional regulation, and epigenetic regulation (Cortázar et al., 2007; Mohan et al., 2010). Additionally, TDG is absent from human cells during S phase due to proteasome-dependent destruction (Hardeland et al., 2007), making it an ideal potential candidate for regulation by $CRL4^{Cdt2}$. However, the requirements for its proteolysis are unknown. Understanding how this destruction is controlled and what other factors contribute to TDG regulation will contribute to a better understanding of DNA repair and epigenetics.

## *1.5* Xenopus *egg extract and* X. laevis *development*

To study DNA replication and repair, the Walter lab uses extracts prepared from unfertilized eggs of the African clawed frog *Xenopus laevis*. In *X. laevis*, females store mature oocytes that are arrested in metaphase II of meiosis by an unknown signal known as the cytostatic factor (CSF) (Masui and Markert, 1971). CSF prevents the destruction of high levels of cyclin B, which cooperates with Cdk1 to prevent separation of chromosomes during metaphase (Wu and Kornbluth, 2008). Female frogs lay CSF-arrested eggs that are then fertilized by sperm secreted by male frogs. Fertilization by sperm triggers cyclin B destruction through a signaling cascade triggered by calcium release from the smooth endoplasmic reticulum (Liu et al., 2007). After the egg transitions from metaphase II into interphase and the maternal and paternal genomes are united, the zygote enters a series of cell cycles consisting of only DNA replication (S phase) and mitosis (M phase) with no intermittent gap phases (Newport and Kirschner, 1982a). During these early cell stages, DNA licensing takes place at the end of mitosis (Gillespie et al., 2012; Li and Blow, 2004; McGarry and Kirschner, 1998). The total length of each cell cycle is only 35

minutes on average (Newport and Kirschner, 1982a). During this time, cells do not grow; the volume of the embryo remains the same size as the pre-fertilized egg, about 1 μL in volume (Arias and Walter, 2004).

The stages of *Xenopus* development have been well characterized by Nieuwkoop and Faber (Bowes et al., 2007; Faber and Nieuwkoop, 1994). Cell divisions are synchronous until the midblastula transition after the 12$^{th}$ division (Nieuwkoop and Faber Stage 8.5 (Faber and Nieuwkoop, 1994)), when progression through S phase slows, and the cell cycle expands to include G1 and G2 phases (Bowes et al., 2007; Newport and Kirschner, 1982a). The timing of these early cleavage divisions is controlled by cyclin fluctuations throughout the cell cycle, and embryos are insensitive to checkpoint-dependent cell cycle slowing in response to DNA damage (Clute and Masui, 1997; Conn et al., 2004; Hensey and Gautier, 1998; Kappas et al., 2000; Kimelman et al., 1987). High numbers of replication forks likely contribute to the embryo's ability to progress through S phase so rapidly (Walter and Newport, 1997). After the early cleavage divisions, the blastula stage (stages 7-9) spans the mid-blastula transition and marks the start of zygotic transcription (Newport and Kirschner, 1982b). Next, during the gastrula stage (stages 10-12), cells migrate (Davidson et al., 2002; Solnica-Krezel and Sepich, 2012; Winklbauer et al., 1996), apoptosis becomes part of the developmental program (Greenwood and Gautier, 2005; Hensey and Gautier, 1998), and regulated gene expression begins (Bouvet et al., 1994; Rupp and Weintraub, 1991; Veenstra et al., 1999; Wormington and Brown, 1983). Neurulation (stages 13-21), during which the nervous system develops and cell differentiation continues, follows gastrulation. After neurulation, embryos undergo further tissue differentiation, organogenesis, and growth during the tailbud stage (stages 22-44) and tadpole stages (Bowes et al., 2007). Tadpoles then proceed through metamorphosis to become frogs.

By crushing unfertilized eggs via centrifugation, we can stimulate calcium release and simulate

fertilization. Cyclin levels in the resulting crude extract fluctuate like they would in fertilized eggs. When provided with demembranated sperm chromatin, *Xenopus* egg extract supports the assembly of nuclei around DNA, origin licensing, and a single round of genome replication (Blow and Laskey, 1988; Newport, 1987). After S phase is complete, extracts can progress into mitosis, segregate their chromosomes into distinct nuclei, and repeat the complete cycle up to five times (Arias and Walter, 2004). DNA damage and fork stalling can also trigger a checkpoint response and DNA repair in extract (Dasso and Newport, 1990; Kumagai and Dunphy, 1999; Kumagai et al., 1998). To specifically study S phase-specific processes, progression into mitosis can be blocked through the addition of the translation inhibitor cycloheximide to inhibit the production of cyclin B and other proteins. Therefore, S phase *Xenopus* egg extract behaves like an *in vivo* system with the advantages of an *in vitro* system.

We can further process crude S phase extract into other fractions that serve different purposes:

> Low speed supernatant (LSS) (Blow, 1993): Because crude S phase extract contains mitochondria, it cannot be frozen and subsequently thawed; mitochondria would lose their structural integrity and trigger apoptosis of the extract. To make LSS, we remove the mitochondria and other contaminants from crude S phase extract via centrifugation, allowing the production of large quantities of extract at one time for future use. Importantly, LSS still contains the membranes necessary to form nuclei. LSS behaves similarly to crude S phase extract; it supports the assembly of nuclei around chromosomal DNA, origin licensing, a single round of DNA replication, DNA repair, checkpoint signaling, and $CRL4^{Cdt2}$ function. DNA replication depends on nuclear envelope formation; however, LSS and crude S phase extract will not support replication of small DNA molecules like plasmids because these DNA templates do not support the formation of nuclei.

High speed supernatant (HSS) (Hua et al., 1997; Newport, 1987): In addition to removing the mitochondria from crude S phase extract, we can also remove the membranes and other organelles via centrifugation to generate HSS. HSS supports PCNA-dependent repair of damaged DNA templates and licensing of sperm chromatin but does not contain the membranes to form nuclei or support chromosomal replication. This allows us to separate licensing from replication through the use of different extracts, much like somatic cells separate these events into G1 and S phase.

Nucleoplasmic extract (NPE) (Lebofsky et al., 2009; Walter et al., 1998): After adding sperm chromatin to crude S phase extract, nuclei assemble around the DNA. From this crude S phase extract nuclear assembly reaction, we can separate the nucleoplasm, containing high levels of S phase-specific Cdk activity, from the cytoplasm to generate NPE. Because the nucleoplasm contains high levels of S phase-specific Cdks, it triggers origin firing of licensed DNA. However, high levels of geminin in NPE are incompatible with pre-RC formation, so only one round of replication is permitted. NPE mimics a nuclear environment in the absence of a physical nucleus, allowing the replication of plasmids. HSS or LSS must first license DNA before the addition of NPE in order for replication to occur. The same environment (a mixture of HSS and NPE) also supports DNA repair, checkpoint signaling, and destruction of $CRL4^{Cdt2}$ substrates.

This extract-based system has many advantages over both cell-based systems and standard reconstituted systems. Extracts lack a cellular membrane that restricts the entry and exit of specific proteins. Removal of proteins in cells requires lengthy techniques like siRNA or genetic manipulations and can, in some cases, only result in partial removal of the target protein. In *Xenopus* egg extract, we can remove proteins through immunodepletion in just a few hours before initiating replication or repair, so no indirect effects of the removal accumulate over time and mask events of interest. Because most DNA replication proteins are essential, their removal

has pleiotropic detrimental effects over multiple cell cycles that can be unrelated to their specific functions. Additionally, adding mutant proteins to extracts has advantages over expressing proteins in cells, such as providing the capacity to ensure specific protein concentrations directly. Finally, *Xenopus* egg extract will replicate or repair almost any DNA template, including plasmids engineered to contain site-specific damage. Nonetheless, compared to traditional *in vitro* systems, *Xenopus* egg extract has the same advantage as cells: extract will recapitulate all the events of replication and repair without requiring the characterization and purification of every protein and molecule involved. Because many of the DNA replication, repair, and checkpoint signaling pathways involve so many players, many of which have not been identified, reconstituted systems are not yet an option to study these pathways on a molecular level. Consequently, *Xenopus* egg extract allows us to study DNA replication and repair on a biochemical and molecular level not possible in tissue culture or in *in vitro* systems based on purified recombinant proteins.

# Chapter 2

# Thymine DNA glycosylase is a $CRL4^{Cdt2}$ substrate

## *2.1 Introduction*

Thymine DNA glycosylase (TDG), a base excision repair (BER) protein, has taken the spotlight in the fields of DNA repair and epigenetics in recent years due to its potential role in DNA demethylation. DNA methylation is an epigenetic mark that regulates gene expression. Though we have a general understanding of how DNA methylation occurs, DNA demethylation remains enigmatic in animals (He et al., 2011a). BER is responsible for DNA demethylation in plants and is proposed to be part of a two-step pathway for demethylation in animals in which 5mC-modifying enzymes first process 5mC to another moiety that can then be removed through BER, generating unmethylated cytosine as a result (Chen and Riggs, 2011; Dalton and Bellacosa, 2012; Fritz and Papavasiliou, 2010; Nabel and Kohli, 2011; Niehrs, 2009; Wu and Zhang, 2010).

In current models, TDG is the most likely candidate for the BER glycosylase involved in this two-step pathway. TDG and other BER proteins cooperate with AID and APOBEC1, deaminases that process 5mC to thymine, which TDG and downstream BER proteins can convert into an unmodified cytosine (Cortellino et al., 2011; Morgan et al., 2004; Popp et al., 2010; Rai et al., 2008). Additionally, TDG functionally interacts with Tet proteins, enzymes that iteratively oxidize 5mC to 5hmC, 5fC, and 5caC. TDG then acts on the resulting 5fC and 5caC, ultimately leading to cytosine demethylation (Guo et al., 2011; Hashimoto et al., 2012; He et al., 2011b; Inoue and Zhang, 2011; Inoue et al., 2011; Iqbal et al., 2011; Ito et al., 2011; Maiti and Drohat, 2011; Wossidlo et al., 2011; Zhang et al., 2012). A more detailed discussion of these models is provided in Chapter 1.

TDG is essential for development. Mice lacking TDG die at embryonic day 11.5. Leading up to their death, the fetal mice suffer from aberrant DNA methylation due to a failure to demethylate gene promoters that must be activated during tissue differentiation (Cortázar et al., 2011; Cortellino et al., 2011). This observation further supports the hypothesis that TDG is important

for DNA demethylation. Moreover, a recent study showed that targeting TDG directly to specific genes results in demethylation of these gene promoters (Gregory et al., 2012), suggesting that localization of TDG is sufficient for DNA demethylation.

Understanding the regulation of a protein involved in so many DNA transactions, ranging from BER to regulation of epigenetic inheritance, is of interest to cell and molecular biologists of all specialties. During BER, TDG is product-inhibited due to virtually irreversible binding to an abasic site; its removal depends on SUMOylation of the C-terminus, which decreases TDG's affinity for DNA (Baba et al., 2005; Hardeland et al., 2002; Smet-Nocca et al., 2011). Additionally, TDG is absent from human cells during S phase due to proteasome-dependent destruction after ubiquitylation by an unknown ubiquitin ligase (Hardeland et al., 2007). The authors who identified this S phase-specific destruction proposed that during DNA replication, the TDG-abasic site interaction could be too strong for the replication fork to overcome, leading to fork stalling during S phase (Hardeland et al., 2007). In support of this model, massive overexpression of TDG causes an accumulation of cells in S phase (Hardeland et al., 2007).

$CRL4^{Cdt2}$ negatively regulates many proteins whose presence in S phase is detrimental (Abbas and Dutta, 2011; Havens and Walter, 2011). $CRL4^{Cdt2}$ targets its substrates for destruction specifically during DNA replication or following DNA damage. Fewer than ten $CRL4^{Cdt2}$ substrates are known, but all have a role in DNA replication or repair (Havens and Walter, 2011). The Walter lab has characterized the $CRL4^{Cdt2}$ degron as well a role for PCNA in $CRL4^{Cdt2}$ recruitment to the substrate. $CRL4^{Cdt2}$ substrates must bind to DNA-associated PCNA through their PCNA interacting peptide (PIP) boxes (Figure 1.2) to function as degradation targets (Arias and Walter, 2006). Ubiquitin ligase recruitment involves contact between Cdt2 and residues on both PCNA and the substrate (Havens and Walter, 2009; Havens et al., 2012). However, $CRL4^{Cdt2}$ does not target all PIP box-containing proteins for destruction. It accomplishes

substrate specificity for a small subset of PCNA-binding proteins through additional degron features: all known $CRL4^{Cdt2}$ substrates contain a basic residue four amino acids downstream of the PIP box (the B+4 position relative to the PIP box), and most also contain a threonine and aspartate at positions 5 and 6 of the PIP box (a "TD motif") (Havens and Walter, 2009), while stable PCNA interacting proteins do not (Chapter 1, Figure 1.2). These features create a "PIP degron" necessary for targeting by $CRL4^{Cdt2}$.

Given our interest in $CRL4^{Cdt2}$-dependent proteolysis, I was intrigued by the cell cycle profile of TDG. The ubiquitin ligase responsible for S phase-specific proteolysis of TDG (Hardeland et al., 2007) has not been identified. Here, I report that $CRL4^{Cdt2}$ regulates TDG, which contains a PIP degron in its N-terminal regulatory region. However, wildtype TDG is destroyed with slower kinetics than the well-characterized $CRL4^{Cdt2}$ substrate Cdt1 in response to damage and not at all during normal replication in *Xenopus* egg extract. Improving TDG's PIP degron expedites TDG destruction in response to damage and triggers its destruction during DNA replication, suggesting that it is a weaker $CRL4^{Cdt2}$ substrate due to decreased PCNA binding compared to other substrates. In developing embryos, non-degradable TDG is regulated similarly to wildtype TDG during the early cleavage divisions but accumulates during gastrulation and beyond, suggesting that $CRL4^{Cdt2}$-dependent destruction of TDG changes during development. Nevertheless, I was unable to identify replication or developmental defects in response to a non-degradable TDG mutant and anticipate that other redundant mechanisms restrict the access of TDG to DNA, thereby masking the effects of non-degradable TDG.

## *2.2 Materials and Methods*

### 2.2.1 Preparation of testes and sperm chromatin from *X. laevis* males

The methods for isolation of testes and production of sperm chromatin from post-mortem male frogs was previously reported (Lebofsky et al., 2009; Tutter and Walter, 2006). Male frogs were

sacrificed, and their testes were isolated through dissection. For subsequent fertilization, testes were stored in 1x Marc's Modified Ringer Solution (MMR; 100 mM NaCl, 2 mM KCl, 1 mM $MgCl_2$, 2 mM $CaCl_2$, 2 mM $NaHCO_3$, 5 mM HEPES, pH 7.8) containing 50 µg/mL gentamycin. To isolate sperm chromatin, testes were harvested and immediately minced with a razorblade and diluted in Buffer X (80 mM KCl, 15 mM NaCl, 5 mM $MgCl_2$, 1 mM EDTA, 10 mM HEPES, pH 7.4) supplemented with 200 mM sucrose. Sperm was released from the testes by repetitive rounds of vortexing and centrifugation at 1000 rpm in an IEC CL2 tabletop centrifuge for 10 seconds. The soluble sperm was removed to a new tube and centrifuged at 1500 rpm to remove any contaminating tissue. Sperm was pelleted by centrifugation in a Sorvall HB6 rotor at 4000 rpm, 4°C for 10 minutes. Sperm was then resuspended in Buffer X supplemented with 2 M sucrose and isolated from red blood cells and other contaminants by centrifugation through a sucrose cushion (Buffer X containing 2.3 M sucrose underlaid with Buffer X containing 2.5 M sucrose) in a TL100 Beckman Ultracentrifuge rotor at 33,000 rpm, 2°C for 25 minutes. Purified sperm was washed in Buffer X containing 200 mM sucrose and recovered through centrifugation in a Sorvall HB6 rotor at 5000 rpm, 4°C for 10 minutes. Sperm was demembranated in Buffer X containing 200 mM sucrose, 10 µg/mL aprotinin, 10 µg/mL leupeptin, 1 mM DTT, and 0.4% Triton X-100 through gentle agitation at 4°C for 30 minutes. Detergent was removed by centrifugation through a sucrose cushion containing Buffer X supplemented with 500 mM sucrose, 10 µg/mL aprotinin, 10 µg/mL leupeptin, 1 mM DTT, and 3% bovine serum albumin (BSA; Sigma) and two washes with sperm chromatin storage buffer (Buffer X containing 200 mM sucrose, 10 µg/mL aprotinin, 10 µg/mL leupeptin, 1 mM DTT, and 3% BSA). The amount of recovered sperm chromatin was determined, and the chromatin was stored in the same buffer in single use aliquots at -80°C.

### 2.2.2 Unfertilized egg collection

Female frogs were injected with 75 international units of pregnant mare serum gonadotropin

(PMSG; Calbiochem) 2-7 days before injection with 625 units of human chorionic gonadotropin (HCG; Sigma) to induce egg laying. 20 frogs were injected to make NPE or 6 frogs were injected to make HSS. For microinjection and development experiments, eggs were freshly squeezed from a single female frog injected for a separate extract preparation.

### 2.2.3 *Xenopus* egg extract

Eggs were harvested 15-22 hours after injection with HCG. Crude S phase extract was prepared as previously reported (Tutter and Walter, 2006) and is described here. To make crude S phase extract as an endpoint or to subsequently make NPE or HSS, eggs were harvested and dejellied with 2.2% cysteine (Spectrum), pH 7.7, washed with ½x MMR, and then washed with egg lysis buffer (ELB; 250 mM sucrose, 2.5 mM $MgCl_2$, 50 mM KCl, 10 mM HEPES, pH 7.7) containing 50 µg/mL cycloheximide and 1 mM DTT. Eggs were packed at 1100 rpm in an IEC CL2 centrifuge at room temperature and supplemented with 5 µg/mL aprotinin and leupeptin and 2.5 µg/mL cytochalasin B. Eggs were activated by crushing via centrifugation at 10,000 rpm, 4°C for 15 minutes in a Sorvall HB6 rotor equilibrated to room temperature. Following centrifugation, crude S phase extract (the middle, soluble layer) was harvested using a syringe. Crude S phase extract was supplemented with 50 µg/mL cycloheximide, 1 mM DTT, 10 µg/mL aprotinin and leupeptin, and 5 µg/mL cytochalasin B and stored on ice until the addition of sperm chromatin or further processing.

HSS was also prepared as previously described (Hua et al., 1997; Newport, 1987; Walter et al., 1998). To make HSS, crude S phase was centrifuged in a TL100 Beckman Ultracentrifuge rotor for 90 minutes at 55,000 rpm, 2°C. The lipids were removed, and HSS was harvested and centrifuged again for 30 minutes at 2°C to remove any additional lipids. HSS was flash frozen in liquid nitrogen and stored at -80° until use.

NPE was prepared according to Walter et al., 1998 with a few modifications. To make NPE,

crude S phase extract was supplemented with 3.3 µg/mL nocodazole, optionally diluted up to 10% by volume with ELB containing cycloheximide and DTT, and further centrifuged for 15 minutes in a Sorvall HB6 rotor at 4°C to eliminate any remaining lipids, mitochondria, and yolk proteins. Lipids were aspirated, and the clean crude S phase extract was decanted, leaving the mitochondria and yolk proteins behind. Clean crude S phase extract was supplemented with 2 mM ATP (Sigma), 20 mM phosphocreatine (Sigma), and 5 µg/mL creatine kinase (Sigma) and equilibrated to room temperature. Demembranated sperm chromatin was added to a final concentration of 4400/µL and allowed to form nuclei. The nuclei assembly reaction was incubated at room temperature and mixed every 10 minutes by gentle inversion for 90 minutes or until nuclei reached an average diameter of 25-30 µm. Nuclei were collected by centrifugation for two minutes at 10,000 rpm, 4°C in a Sorvall HB-6 rotor. The upper nuclear layer was removed and centrifuged at 55,000 rpm (260,000xg) at 2°C in a TL55 swinging bucket rotor in a Beckman TL100 tabletop ultracentrifuge. Any contaminating lipids were removed, and the nucleoplasm was harvested, frozen on liquid nitrogen, and stored at -80°C until use.

Before use, NPE was diluted with 40-60% ELB to elicit optimal DNA replication. All extracts were supplemented with an energy regeneration mix (2 mM ATP, 20 mM phosphocreatine, and 5 µg/mL creatine kinase). HSS and crude S phase extract were also supplemented with nocodazole (3-5 µg/mL), and diluted NPE was supplemented with an additional 10 mM DTT. All experiments in *Xenopus* egg extract described in this chapter were performed in a 2:1 mixture of diluted NPE:HSS, unless otherwise noted. DNA was incubated first in HSS before the addition of diluted NPE for replication assays (described below in section 2.2.9).

### 2.2.4 Egg fertilization

Eggs were fertilized according to published protocols (Sive et al.); Eggs were harvested directly into a petri dish containing 1x MMR. Fertilization was carried out at 18°C. Liquid was removed,

and a small piece of testes that was lightly minced with a razorblade before use was waved over the eggs. Remaining testes chunks were homogenized with a pestle, resuspended in a small volume of 1xMMR, and drizzled over the eggs. The sperm were activated through the addition of an excess of water to the sperm-covered eggs for a final MMR concentration of approximately 0.1x. Fertilized eggs oriented themselves with the animal pole facing upwards. Embryos were dejellied with 2% cysteine, pH 7.8 and washed well with 0.1xMMR approximately 25 minutes after fertilization and stored in 0.1xMMR containing 50 µg/mL gentamycin.

### 2.2.5 *X. laevis* microinjections and developmental experiments

To prepare mRNA for microinjections, pCS2+TDG plasmid (described below) was linearized with a NotI (NEB) overnight digest and purified using a PCR purification kit (Qiagen). mRNA was prepared using the mMessage mMachine Sp6 kit (Ambion) and purified using an RNeasy mini kit (Qiagen). Purified mRNA was ethanol precipitated and diluted in RNase-free water.

200 pg or 1 ng TDG mRNA (10 nL at 2 ng/mL or 10 ng/mL in water) was microinjected into each cell of dejellied Stage 2 embryos in 0.1x MMR+5% Ficoll containing gentamycin (50 µg/mL) at 18°C. Embryos were allowed to heal in 0.1x MMR+5% Ficoll containing gentamycin for 1-2 hours at 18°C and then moved to 0.1x MMR containing gentamycin and stored at 14°C until they reached the midblastula transition (N-F Stage 8.5). Development was then monitored at 23°C. Embryos were staged according to the Nieuwkoop and Faber anatomical stages (Bowes et al., 2007; Faber and Nieuwkoop, 1994). Two microinjected embryos were harvested at the appropriate stages by freezing with liquid nitrogen and storing at -80°C. To monitor endogenous protein levels, pools of 10 uninjected embryos were harvested.

Frozen embryos were lysed in 15 µL embryonic lysis buffer per embryo (250 mM sucrose, 1% NP-40, 10 mM EDTA, 25 mM HEPES, pH 7.5 supplemented with a Roche Complete protease inhibitor tablet) via homogenization with a pipet tip and light vortexing. Yolk proteins were

removed via centrifugation at 10,000xg, 4°C for 10 minutes. The soluble fraction was recleared of yolk proteins with a second spin.

### 2.2.6 Plasmid construction, mutagenesis, and protein purification

A cDNA library was generated from purified *X. laevis* egg mRNA (generated by Dr. Courtney Havens), and the TDG gene was amplified with primers 5'-GCTAAGGATCCATGGAGGCCC AGGACCCAAGC-3' and 5'-TTTTTCTCGAGTCAAGCGTTGCTGCCTCCTTGC-3'. The TDG gene was inserted into a modified pET28 vector that contains a Prescission Protease (GE Healthcare) cleavage site for expression in bacteria or into a pCS2+ vector for mRNA production at BamHI and XhoI (NEB) sites in the respective vector. TDG point mutants were generated with a QuikChange II Site-Directed Mutagenesis Kit (Agilent).

To obtain native, functional TDG, expression of His-tagged TDG was induced in Arctic Express cells (Agilent) with 0.1 mM IPTG for 24 hours at 15°C, and TDG was purified via batch purification. Cells were lysed via sonication in lysis buffer (750 mM NaCl, 20% glycerol, 10 mM β-mercaptoethanol, 10 mM imidazole, 50 mM sodium phosphate, pH 8.0, and protease inhibitors (0.5 mM PMSF; Amresco, 1 mM benzamidine; Sigma, 10 µg/mL aprotinin, 10 µg/mL leupeptin)) and cleared via centrifugation at 30,000g. Soluble His-tagged TDG was bound to Ni-NTA beads (Qiagen) for one hour and washed twice with lysis buffer containing 20 mM imidazole and once with lysis buffer containing 50 mM imidazole (aprotinin and leupeptin were omitted from these wash buffers). To separate the chaperonin protein from TDG, an additional 2 hour wash was performed in an ATP-based chaperonin removal buffer (375 mM NaCl, 20% glycerol, 10 mM β-mercaptoethanol, 10 mM imidazole, 10 mM $MgCl_2$, 150 mM KCl, 5 mM ATP, and 50 mM sodium phosphate, pH 7.6) modified from a previous report on removal of co-purifying chaperonin from Arctic express cells (Joseph and Andreotti, 2008). Finally, the beads were washed again twice with lysis buffer containing 20 mM imidazole and once with lysis buffer

containing 50 mM imidazole before eluting in lysis buffer containing 400 mM imidazole. 10 units of Prescission Protease was added to the native protein eluate, which was then dialyzed overnight in 50 mM NaCl, 50 mM Tris, 10% glycerol, 10 mM β-mercaptoethanol, and 1 mM EDTA, pH 8. TDG was used at a final concentration in extract of 1 ng/μL in this chapter unless otherwise indicated.

To obtain denatured TDG for antibody production, TDG was expressed in BL21 (DE3) cells via induction with 0.25 mM IPTG at 37°C for three hours. Cells were lysed in 7 M urea, 10 M sodium phosphate, and 10 mM Tris, pH 8.0, cleared via centrifugation at 30,000xg, and bound to Ni-NTA beads for one hour. Beads were washed with 8 M urea, 100 mM sodium phosphate, 10 mM Tris, pH 6.3. His-TDG was eluted directly into 2x SDS-PAGE loading buffer (20% glycerol, 6.1% SDS, 125 mM Tris-HCl, pH 6.8, 0.01% Bromophenol Blue), and separated on an SDS-PAGE gel. Protein was stained with Gelcode Blue (Thermo Scientific) and electroeluted into SDS-PAGE running buffer (0.1% SDS, 250 mM glycine, 25 mM Tris, pH 8.3). The resulting purified TDG was >99% pure as measured by Coomassie staining.

Human TDG was donated by Dr. Primo Schär from the University of Basel in Switzerland. p21 peptides (previously described by (Arias and Walter, 2006)) were synthesized commercially. Wildtype p21 peptide (CKRRQTSMTDFYHSKRRAIAS) was synthesized by Invitrogen (Carlsbad, CA); mutant peptide (CKRRATSATDAAHSKRRAIAS) was synthesized by the Biopolymers Laboratory (Harvard Medical School, Boston, MA).

### 2.2.7 Immunological methods

Denatured His-TDG was sent to Pocono Rabbit Farm and Laboratory (Canadensis, PA) for antibody production in two rabbits. Serum from rabbit #197 was used for all TDG immunoblotting at a concentration of 1:5000-1:2500. Serum from rabbit #198 was used for immunoprecipitation. Antibodies to *Xenopus* Cdt1 (Arias and Walter, 2005), Cdt2 (Jin et al.,

2006), Orc2 (Walter et al., 1998), MCM7 (Walter and Newport, 2000), Rcc1 (Dasso and Newport, 1990), and RPA (Walter and Newport, 2000) were generated and described previously. Commercial antibodies were used to blot for PCNA (Santa Cruz), human Cdt1 (Bethyl), human Cdt2 (Bethyl), phospho-Chk1 (S345, Cell Signaling), and total Chk1 (Cell Signaling). Antibodies to human TDG were donated by Dr. Primo Schär.

Immunodepletions were performed using a 3:1 ratio of antibody serum prebound to Protein A Sepharose Fastflow resin (Amersham Biosciences). Antibody-bound resin was used at a 1:5 ratio to egg extract. To deplete Cdt2, either one or two one-hour rounds were performed, both leading to the same results shown.

To immunoprecipitate TDG from extract, 1 µL crude serum (rabbit #198) was added to 10 µL Protein A Sepharose Fastflow resin and incubated at 4°C overnight to bind the TDG antibody to the beads. The beads were washed five times with PBS and two times with ELB containing 500 mM NaCl. Extract reactions (HSS/NPE) were diluted 6-fold in ELB and applied to the beads for 1 hour at 4°C. Bead-bound proteins were washed five times in ELB containing 0.1% Triton X-100 and eluted directly into 2x sample buffer.

### 2.2.8 Generation of MMS-damaged DNA and DNA recovery from extract

Methylated plasmid or linear DNA was generated as previously described (Stokes, 2003): DNA was diluted to a concentration of 250 ng/µL in TE buffer (10 mM Tris, 1 mM EDTA, pH 8.0) and then diluted 1:1 with 1xM9 salts (Sambrook and Russell, 2001) containing 0.1 mM $MgSO_4$. MMS was added to a final concentration of 450 mM and incubated with DNA at 30°C for 30 minutes. MMS was removed through ethanol precipitation. 5-15 ng of MMS-damaged plasmid DNA was used to trigger TDG destruction in extract. Unless otherwise indicated, MMS-damaged plasmid was used at 10 ng/µL in this chapter.

To make linear, bead-bound, MMS-damaged DNA, PCR was used to generate a double-stranded 1kb linear DNA fragment that was biotinylated on both ends. The DNA fragment was purified using a PCR purification kit (QIAGEN) before MMS treatment. The resulting damaged, biotinylated fragment was bound to M-280 streptavidin Dynabeads (100 ng DNA per 10 mg Dynabeads; Invitrogen) by incubating the DNA and beads for 2-4 hours at 23°C in binding buffer (10 mM Tris 8.0, 100 mM NaCl, and 1 mM EDTA) and subsequently incubating overnight at 4°C. Unbound DNA was removed, and an excess of streptavidin was added to block any free biotin ends before washing the bead-bound DNA three times with the binding buffer. Beads were stored in binding buffer at the original bead volume. Approximately fifty percent of the DNA bound to beads. Buffer was removed before the addition of extract.

To recover immobilized MMS-damaged DNA from extract, samples were resuspended in ELB containing 50-150 mM KCl and 0.6% Triton X-100 and then layered over an ELB cushion containing 500 mM sucrose and 50-150 mM KCl. Following centrifugation through the sucrose cushion, the beads were washed with resuspension buffer.

### 2.2.9 DNA replication assays and nascent strand analysis in *Xenopus* egg extract (NPE)

DNA was licensed in HSS before the addition of diluted NPE containing trace amounts of α-$^{32}$P-dATP. Where indicated, licensing was inhibited with 400 nM geminin. NPE mimics the nuclear environment, allowing origin firing. At the desired time points, replication was stopped in replication stop buffer (8 mM EDTA, 10% Ficoll, 5% SDS, 0.2% bromophenol blue, 80 mM Tris, pH 8) and subsequently treated with proteinase K (1 mg/mL; Roche) before separation on a 0.8% agarose gel. Gels were dried on filter paper and visualized with a phosphorimager. Replication was monitored by the incorporation of α-$^{32}$P-dATP and normalized to total radioactive signal in each sample. The concentration of dATP in each extract preparation was calculated separately by measuring the effect of adding a known concentration of unlabeled

dATP. Total replication was quantified based on the normalized α-$^{32}$P-dATP incorporation, concentration of dATP in extract, and concentration of total parental DNA in each sample.

For nascent strand analysis, the plasmid products were digested with PvuII to generate a 448nt parent-strand fragment hybridized to nascent strands of varying lengths. Samples were diluted in formamide loading dye (Ambion) and separated on 42 cm long, 7% polyacrylamide sequencing gels (7 M urea, 0.8x glycerol-tolerant buffer (USB), the desired concentration of acrylamide (from a 40% Rapid Gel XL concentrate; USB), and APS and TEMED for polymerization). Gels were dried on filter paper and visualized with a phosphorimager. Sequencing ladders were prepared with the Cycle Sequencing Kit (USB) and included on each gel to measure band sizes with nucleotide resolution.

### 2.2.10 TDG activity assays

Base release assays were used to measure TDG activity on a short oligo template and on a plasmid containing a G-T mispair (Figure 2.1) (Hardeland et al., 2000; He et al., 2011b; Maiti and Drohat, 2011). 29mer double-stranded oligonucleotides were prepared by end labeling the strand containing the base to be hydrolyzed by TDG (5'-CCGCTGAGGGATATNGAATTCC TGCAGGC-3', where N=C, T, or U) with γ-$^{32}$P-ATP and annealing an unlabeled complementary strand (5'-GCCTGCAGGAATTCGATATCCCTCAGCGG-3') by heating the mixture to 85°C and allowing it to cool slowly to 23°C. Polynucleotide kinase (PNK; NEB) was used to attach the radiolabeled phosphate in PNK buffer (10 mM $MgCl_2$, 5 mM DTT, 70 mM Tris pH 7.6). The DNA strands were annealed in the same buffer. Free label was removed with G-50 Probequant columns (GE Healthcare).

Plasmid containing a single G-T mispair was generated in the Takahashi lab at Osaka University in Japan by annealing a mispaired oligonucleotide to a single-stranded plasmid and performing strand extension and ligation.

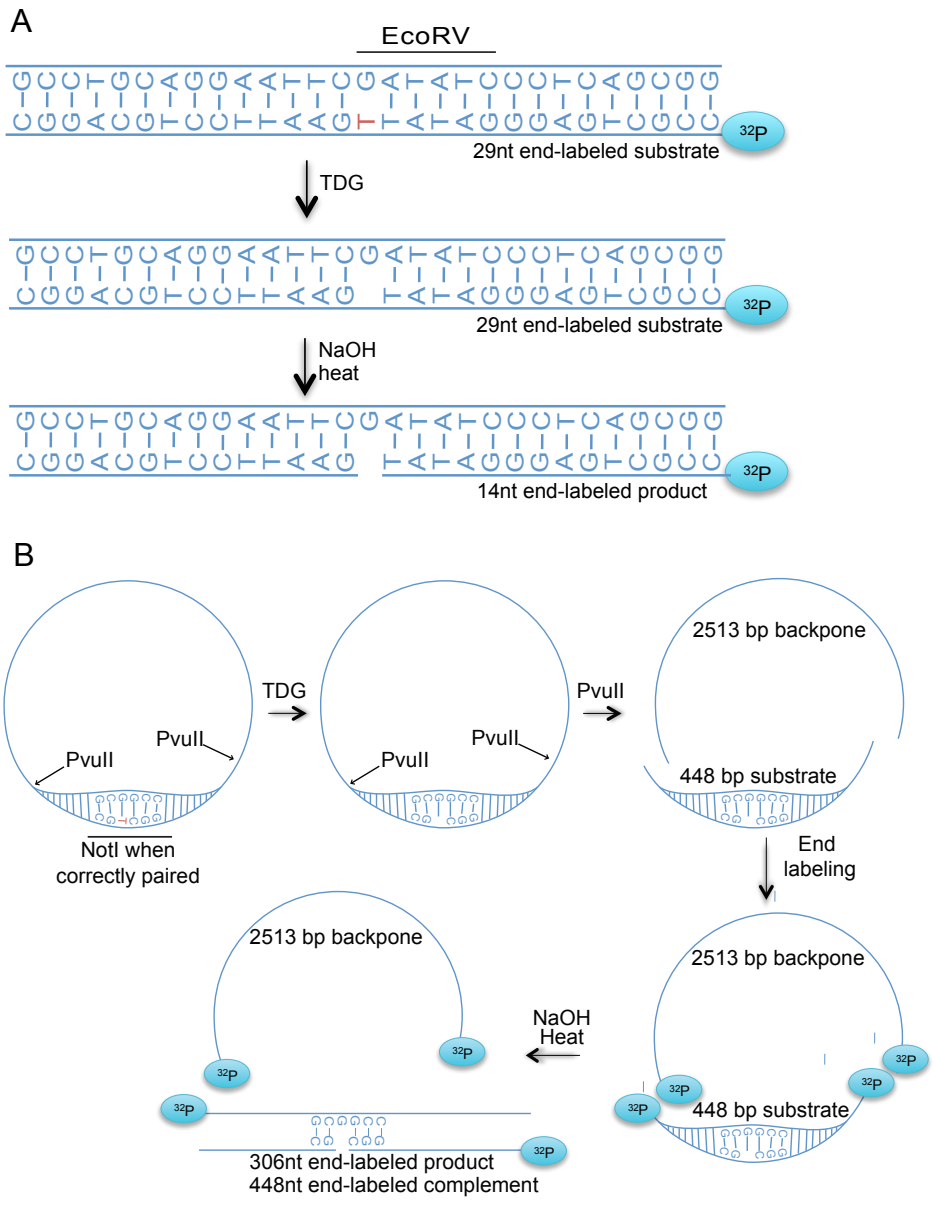
TDG was incubated with either double-stranded 29mer or plasmid in extract, ELB, or activity buffer (20 mM NaCl, 0.5 mM EDTA, 1 mM DTT, 50 mM Tris, pH 7.5, and 10 μg/μL aprotinin and leupeptin, along with trace amounts of β-mercaptoethanol and aprotinin and leupeptin contributed by the protein itself). Activity assays were incubated for 40 minutes at 23°C unless otherwise indicated. For plasmid activity assays, the reaction was then stopped in stop solution (25 mM EDTA, 0.5% SDS, 50 mM Tris, pH 7.5). Proteinase K was added to remove TDG from the plasmid, and a phenol chloroform extraction was performed to isolate the DNA. If the activity assay took place in extract, samples in stop solution were incubated with RNase (80 ng/μL; Roche) for 30 minutes at 37°C before the addition of Proteinase K to eliminate RNA from the reaction. The plasmid was concentrated via ethanol precipitation and subsequently digested with PvuII (NEB) to release a 443 base pair fragment for subsequent treatment. For activity assays on the short oligonucleotides, the DNA was processed directly without any further digests or purification steps. The TDG-generated abasic sites were cleaved after the addition of 90 mM NaOH and heating to 95° for three minutes. An equimolar volume of acetic acid was added to neutralize the reaction, and samples were diluted in formamide loading dye (Ambion). Samples were reheated to 75°C and loaded immediately on a small urea-PAGE gel (7 M urea, 0.8x glycerol-tolerant buffer (USB), 7% polyacrylamide (from a 40% Rapid Gel XL concentrate; USB), and APS and TEMED for polymerization) that had been pre-run for 20 minutes at 200V in 0.8x glycerol-tolerant buffer. Gels were run at 200V until the size range of interest reached the middle of the gel. TDG activity was visualized by a change in the size of the end-labeled DNA strand upon exposure with a phosphorimager.

### 2.2.11 Human cell tissue culture and siRNA

HeLa cells were maintained in DMEM (Invitrogen). Reverse siRNA methods were used to

**Figure 2.1: TDG activity assay scheme. A)** A base release assay using short oligos as a template is shown. In a base release assay, following TDG activity and treatment with NaOH, the substrate and product are separated on a 20% urea-PAGE gel. TDG activity results in the generation of a 14nt product, whereas an EcoRV digest of the correctly paired oligos results in a 12nt product. **B)** TDG activity assay on a plasmid containing a G-T mispair. After TDG activity, the plasmid is repurified with a phenol-chloroform extraction and ethanol precipitation. The plasmid is then digested with PvuII to generate a short fragment that is treated like the oligos in (A). TDG activity results in the generation of a single end-labeled product of 306nt along with an uncleaved end-labeled complement of 448nt. A NotI digest of the correctly paired plasmid results in complete digest to two end-labeled fragments of 306nt and 142nt. Substrate and product are separated on a 5% urea-PAGE gel.

**Figure 2.1** (continued)



eliminate Cdt2 from HeLa cells: siRNA was combined with RNAi Max transfection reagent activated in Optimem (Invitrogen) and mixed with cells in suspension. siRNA-containing media was removed after the cells adhered to the plate, and cells were treated for 72 hours before any experimentation. siRNA was purchased from Invitrogen (siRNA1:DTLHSS182090, siRNA2:DTLHSS122293, siRNA3:DTLHSS122295). To damage cells with UV irradiation, media was removed, and cells were placed in a warmed up UV Stratalinker set to 50J/$m^2$. To arrest cells at the start of S phase, a double thymidine block was performed with media containing 2.5 mM thymidine. Cells were treated with thymidine for 16-18 hours, released for 8 hours, and treated again for 16 hours before releasing into S phase. The timing of the thymidine block was coordinated with the timing of siRNA treatment.

For cell cycle analysis, cells were fixed in 70% cold ethanol in PBS for at least 4 hours at 20°C. Ethanol was removed by washing with PBS washing, and cells were resuspended in PBS containing 15 µg/mL propidium iodide and 200 µg/mL RNase. DNA content was analyzed with a FACS machine (BD).

## *2.3 Results*

### 2.3.1 Identification of a putative PIP degron in TDG's N-terminus

Given that TDG levels specifically drop in S phase due to proteasome-dependent destruction (Hardeland et al., 2007), we looked for and identified a putative PIP degron in the N-terminus of TDG (Figure 2.2). Though TDG lacks one of the two conserved aromatic residues in its putative degron, it contains the other conserved elements of a PIP box (a Q or N at position 1, a hydrophobic residue at position 4, and an aromatic residue at position 8), a basic residue four amino acids downstream of the PIP box (B+4 residue), and a threonine and aspartate (TD motif) at positions 5 and 6 (Figure 2.2). These features create a "PIP degron" that is critical for full destruction of all known $CRL4^{Cdt2}$ substrates (Havens and Walter, 2009).

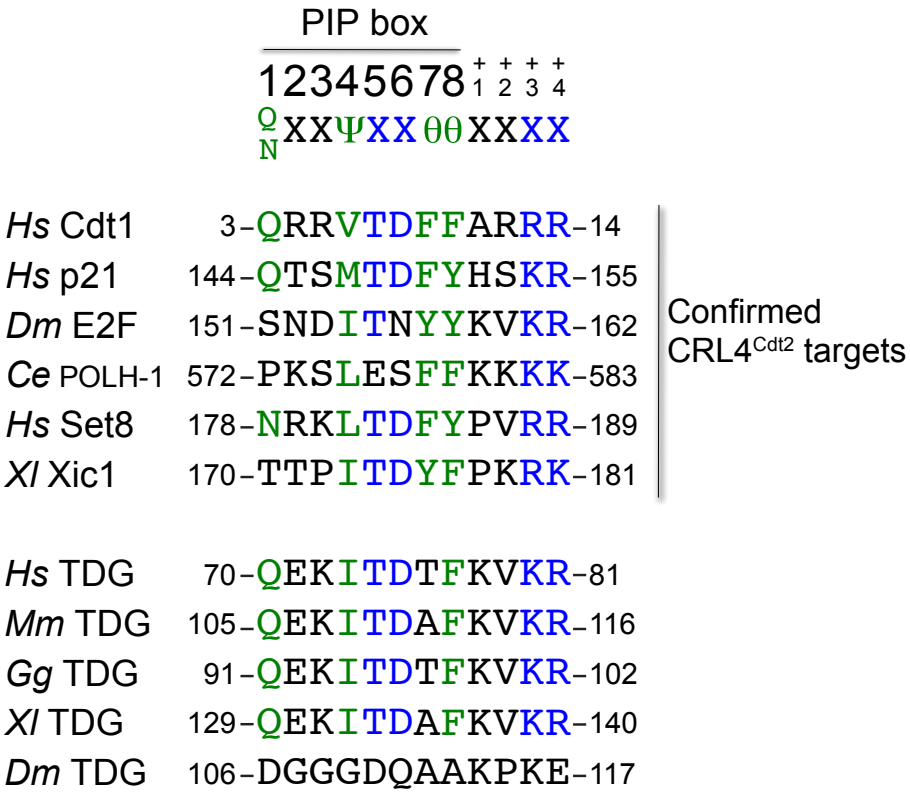



**Figure 2.2: The putative PIP degron of TDG.** The PIP degrons of confirmed CRL4$^{Cdt2}$ substrates are shown above TDG's putative PIP degron in human, mouse, chicken, and frog. TDG of lower eukaryotes, including *Drosophila*, does not contain a PIP box in the same region of the protein or elsewhere. Ψ represents any hydrophobic residue (I/L/V/M), and θ represents any aromatic residue (Y/F/W). Residues shown in green are part of the conserved PIP box, and residues shown in blue additionally contribute to the "PIP degron."

### 2.3.2 TDG is destroyed in a Cdt2-dependent manner in human cells

So far, all *bona fide* vertebrate $CRL4^{Cdt2}$ substrates are destroyed in both *Xenopus* egg extract and human cells during DNA replication and repair. I therefore monitored TDG levels following DNA damage in human cells. TDG levels decreased in HeLa cells and in T98G cells after UV irradiation with $50J/m^2$ UV-C light (Figure 2.3A and data not shown). In addition, siRNA against Cdt2 (Figure 2.3B) stabilized TDG after UV damage (Figure 2.3C, compare lanes 4-12 with lanes 1-3) and during S phase in HeLa cells (Figure 2.3D, compare lanes 4-8 with 10-14), suggesting that human TDG is a $CRL4^{Cdt2}$ substrate. siRNA against Cdt2 also stabilized Cdt1 (Figure 2.3C, D). However, siCdt2 also caused substantial G2/M arrest (Figure 2.3E), so I cannot rule out an indirect effect of siCdt2 on TDG levels due to its effect on the cell cycle from these data alone.

### 2.3.3 Purification and characterization of *Xenopus* TDG

To characterize TDG as a $CRL4^{Cdt2}$ substrate on a molecular level, I moved to the *Xenopus* egg extract system. I purified recombinant *Xenopus* TDG from bacteria (Figure 2.4A) and used it to immunize rabbits. The resulting *Xenopus* TDG antiserum detects recombinant *Xenopus* TDG (Figure 2.4B, lanes 1-5).

It was difficult to detect TDG in unfertilized egg extract (Figure 2.4B, C), leading us to suspect that TDG might be developmentally regulated. A high throughput analysis of absolute mRNA levels during *X. laevis* development previously showed that TDG mRNA levels are high compared to many developmentally-regulated transcripts (Yanai et al., 2011). TDG mRNA levels reported in this published study are shown compared to Pax6 mRNA (Yanai et al., 2011), which is not expressed until the midblastula transition, in Figure 2.4D. I monitored TDG levels during the course of *X. laevis* development and found that they were very low (less than 5 nM) during the early cleavage divisions and blastula stage despite fairly high TDG mRNA levels

**Figure 2.3: Human TDG is destroyed during DNA repair and S phase in a Cdt2-dependent manner in HeLa cells. A)** TDG is destroyed following exposure to 50J/$m^2$ UV irradiation. **B)** siRNA against Cdt2 eliminates Cdt2 protein from cells. **C)** siRNA against Cdt2 stabilizes TDG during DNA repair. Three independent siRNAs were used. **D)** siRNA against Cdt2 leads to accumulation of TDG during S phase. Cells were arrested at the transition to S phase with a double thymidine block and then released into S phase and monitored for TDG destruction. A number of bands detected by the human TDG antibody accumulate upon siCdt2 treatment. Recombinant human TDG is shown in lane 1, and recombinant TDG after incubation in HSS is shown in lane 2 to indicate the expected size of human TDG. **E)** Removing Cdt2 from HeLa cells leads to a substantial G2 arrest. A fraction of cells from (D) were analyzed by FACS for DNA content. PCNA is shown in all panels as a loading control. The experiments shown in this figure were performed with the assistance of Dr. Malavika Raman of the Harper lab at Harvard Medical School.

**Figure 2.3** (continued)

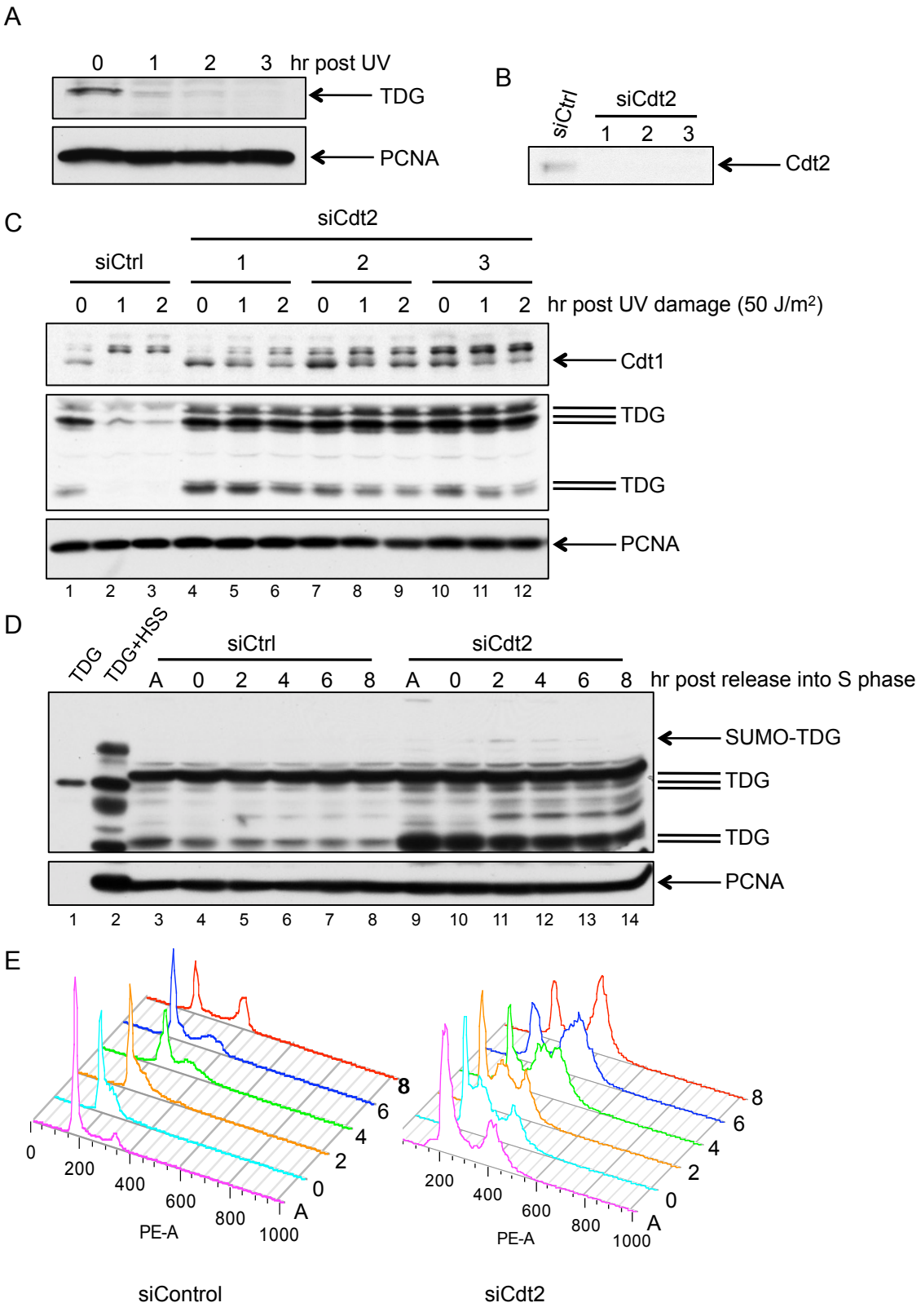
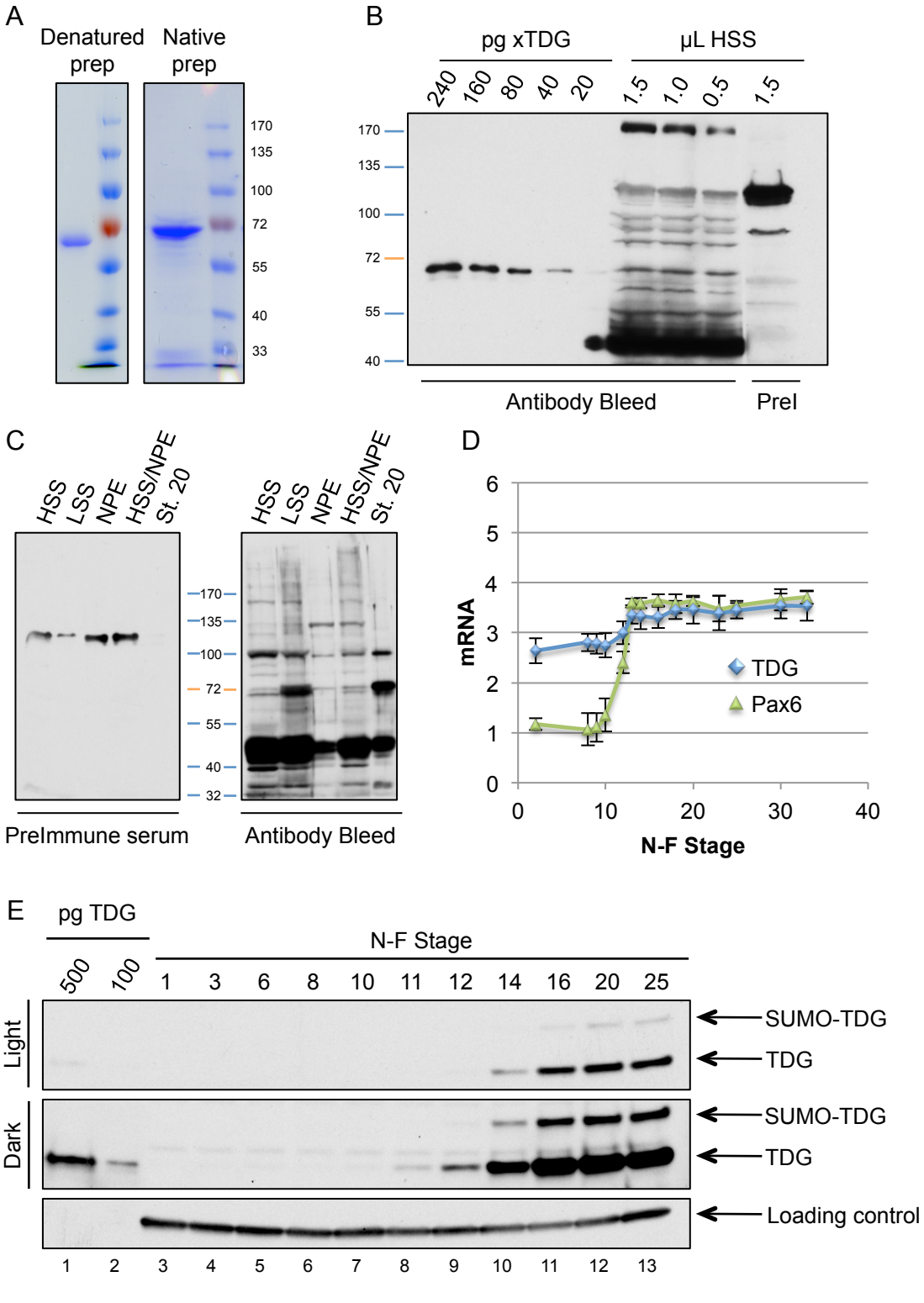


**Figure 2.4: Characterization of *Xenopus* TDG. A)** Denatured TDG was prepared from the insoluble fraction of BL21(DE3) *E. coli* (left) to generate specific antibodies to *Xenopus* TDG, and native TDG was prepared from the soluble fraction of Arctic Express *E. coli*. **B)** Anti-TDG (1:2500) generated from denatured TDG as an antigen detects even low levels of recombinant TDG and endogenous TDG in HSS, while pre-immune serum (1:1000) does not. **C)** Anti-TDG detects low levels of TDG in all extracts used in this work (0.6 µL HSS, 0.6 µL LSS, 0.3 µL NPE, and 0.6 µL diluted NPE/HSS) but detects much higher levels of TDG in lysates from Stage 20 *X. laevis* embryos (0.13 µL). **D)** TDG mRNA levels increase slightly during the course of *X. laevis* development. Pax6 mRNA, which is not substantially expressed until after the mid-blastula transition, is shown for comparison. mRNA (y-axis) is reported as $\log_{10}$ relative concentrations of mRNA abundance over three separate probes for a given gene. Values for (D) were obtained from published mRNA levels measured in three separate batches of fertilized eggs (Yanai et al., 2011). **E)** TDG is barely detectable in early *X. laevis* embryos but dramatically increases during gastrulation and early neurulation. PCNA is shown as a loading control.

**Figure 2.4** (continued)



throughout early development (Figure 2.4D, E). TDG levels increased more than 10-fold during gastrulation and neurulation (>50 nM) (Figure 2.4E), even though TDG mRNA levels did not increase similarly (Figure 2.4D). These results suggest that early embryos regulate TDG at the translational or post-translational level, but that this regulation is lifted later in development.

I tested recombinant *Xenopus* TDG for activity on G-T and G-U mispairs in a base release assay: after TDG hydrolyzed the base from the backbone, sodium hydroxide and heat caused backbone breakage at the abasic site, generating a 14-nucleotide product visible on a denaturing polyacrylamide gel (Figure 2.1). As a control, I digested correctly paired DNA with a restriction enzyme that cleaves near the mispair (Figure 2.5A, lanes 1 and 10). TDG and all mutants expected to retain catalytic activity were active on short oligos containing a G-T mispair (Figure 2.5A, lanes 4-8 and 12-1). As expected, a catalytically inactive mutant ("CI," N173A) failed to cause base hydrolysis (Figure 2.5A, lanes 9 and 18). TDG also acted on G-T and G-U mispairs in the context of a plasmid (Figure 2.5B, lanes 6-7 and Appendix A4, Figure A4.3).

After confirming that recombinant TDG is active, I added it to extract to study its regulation. Recombinant TDG was rapidly modified in extract (Figure 2.5C), and I noticed that a fraction of endogenous TDG was also modified to a higher molecular weight (Figure 2.4B, C, E). I mutated the known SUMOylation site on TDG ("TDG ΔSUMO," K364A) and found that this modification disappeared (Figure 2.5D). SUMOylation decreases TDG's affinity for DNA to enhance dissociation from its catalytic product, an abasic site, during base excision repair (Baba et al., 2005; Hardeland et al., 2002; Smet-Nocca et al., 2011). In agreement, I found that TDG ΔSUMO associated with DNA more stably than wildtype TDG (Figure 2.5D). SUMOylation in extract did not depend on the addition of DNA or the presence of a nuclear environment (Figure 2.5C), suggesting that *Xenopus* egg extract and embryos regulate the access of TDG to DNA via SUMOylation, even in the cytoplasm.

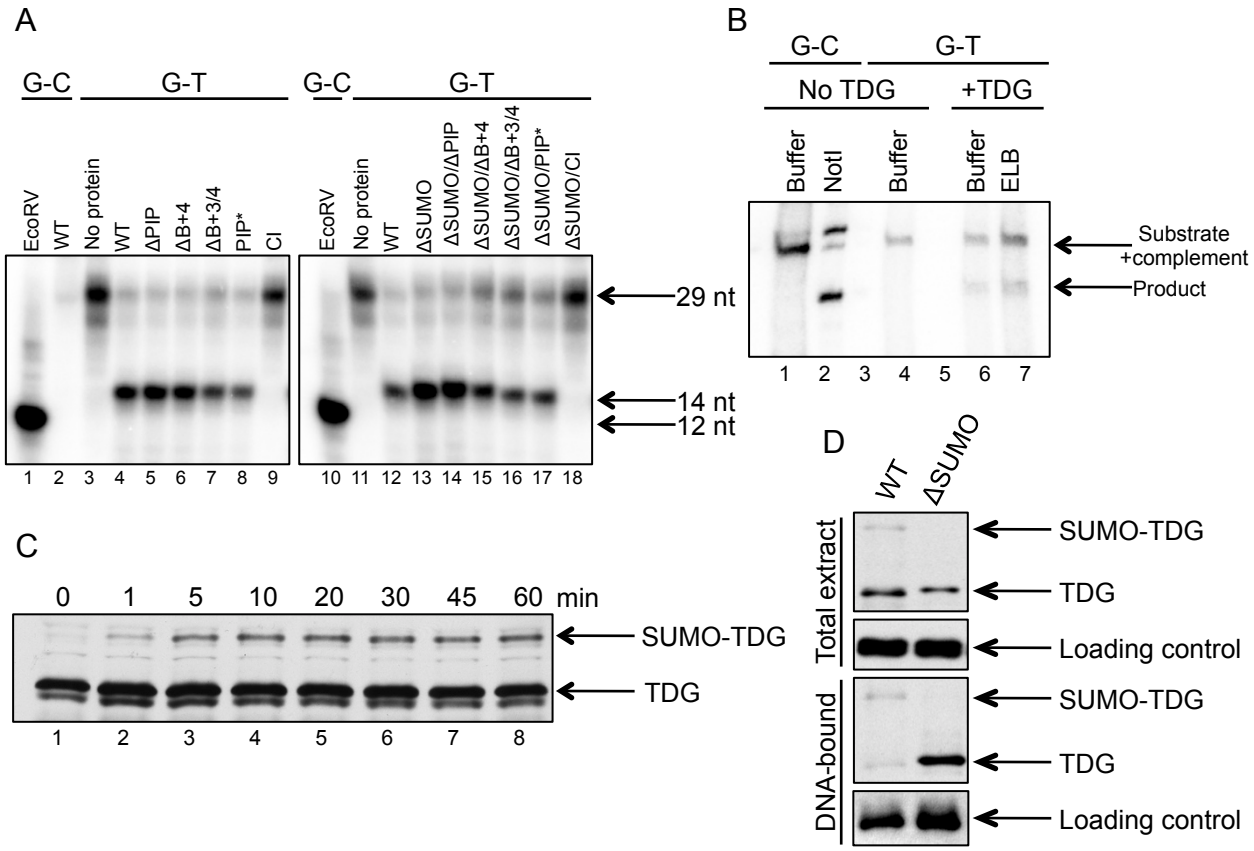



**Figure 2.5: Characterization of recombinant *Xenopus* TDG. A)** Recombinant wildtype TDG and all mutants used in this work are active in a base release assay on the 29mer template (see Figure 2.1) **B)** TDG acts on a G-T mispair in a plasmid. 250 nM TDG ΔSUMO/ΔPIP was used in this reaction. **C)** Recombinant TDG is rapidly modified upon addition to HSS. No DNA was present in this reaction. **D)** Mutation of the SUMOylation site (K364A) eliminates modification of TDG to a higher molecular weight and increases TDG's affinity for DNA. Methyl ubiquitin was included in to allow monoubiquitylation on multiple sites but to prevent polyubiquitylation and destruction. An extract-specific but non-TDG band detected by the TDG antibody was used as a loading control in total extract, and Orc2 was used as a loading control for the DNA-bound fraction.

### 2.3.4 Proteasome- and Cdt2-dependent TDG destruction during DNA repair in *Xenopus* egg extract

To study the regulation of TDG by $CRL4^{Cdt2}$, I added recombinant TDG to HSS/NPE, which contains very low levels of endogenous TDG (Figure 2.4). Known $CRL4^{Cdt2}$ substrates must interact with PCNA on DNA to be targeted for destruction, either during DNA repair or during DNA replication. I therefore asked whether TDG would be destroyed during DNA repair of a DNA template containing methylation damage (but not 5mC) that was generated by treating a plasmid with methyl-methanesulfonate (MMS). Both cells and egg extract repair MMS damage through PCNA-dependent nucleotide excision repair (NER), during which $CRL4^{Cdt2}$ substrates interact with DNA-associated PCNA and are destroyed. I found that an MMS-damaged plasmid triggered TDG destruction, while an otherwise identical undamaged plasmid did not (Figure 2.6A).

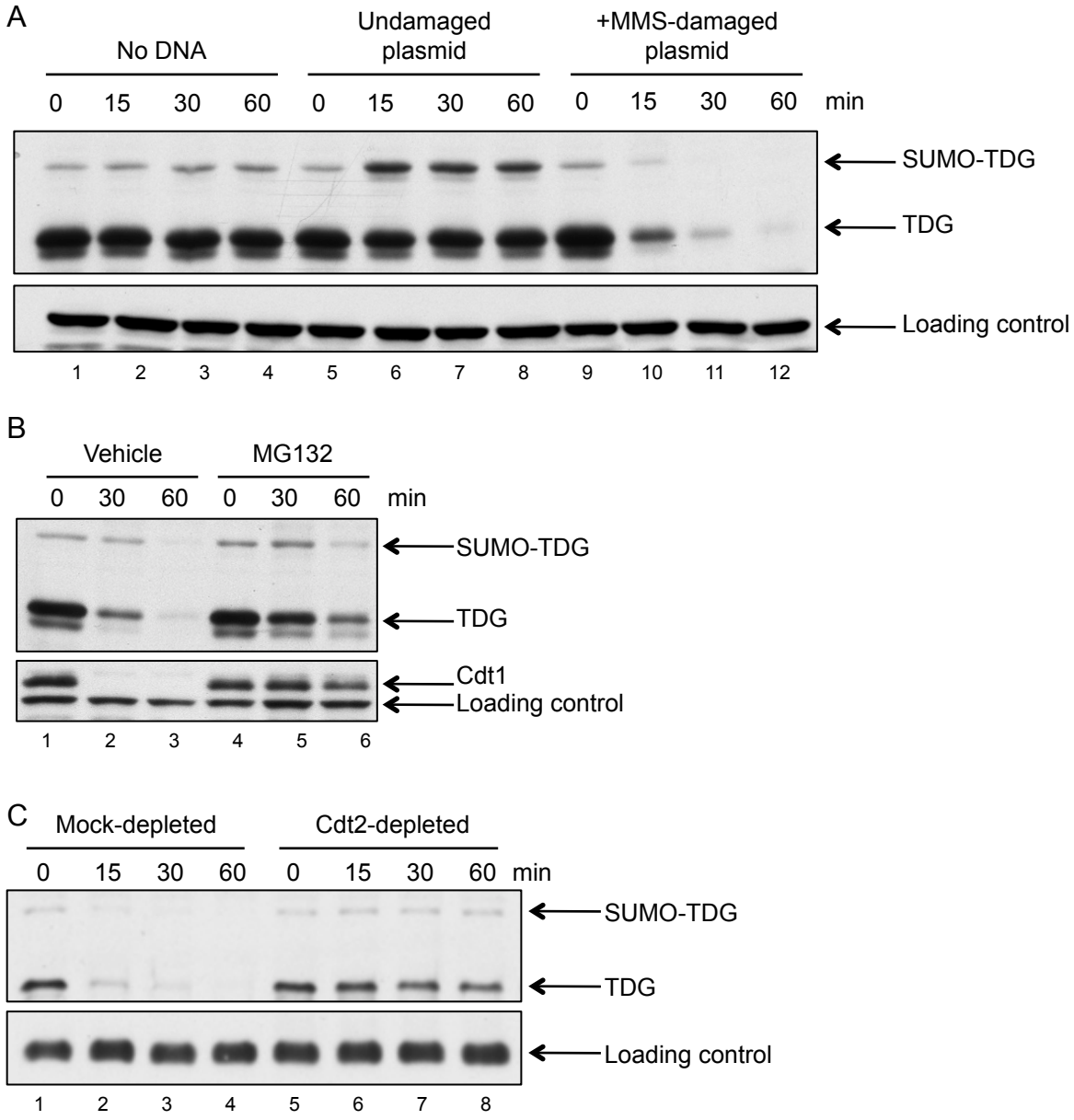
In human cells, TDG destruction during S phase depends on the proteasome (Hardeland et al., 2007). I confirmed that this damage-dependent destruction is also proteasome-dependent by adding the proteasome inhibitor MG132. MG132 impaired TDG destruction in response to damaged DNA (Figure 2.6B). I next tested whether destruction depends on $CRL4^{Cdt2}$ itself. Immunodepletion of Cdt2, the substrate specificity factor for $CRL4^{Cdt2}$, stabilized TDG (Figure 2.6C), suggesting that TDG is a true $CRL4^{Cdt2}$ substrate.

### 2.3.5 The role of SUMOylation in the destruction of TDG

Though TDG destruction depends on Cdt2, DNA damage, and the proteasome, it is much slower than destruction of other known $CRL4^{Cdt2}$ substrates, like Cdt1. Because SUMOylation decreases the affinity of TDG for DNA and TDG ΔSUMO interacts with DNA more stably than wildtype TDG (Figure 2.5C), I hypothesized that SUMOylation might protect TDG from destruction, which is triggered on DNA. If this hypothesis is correct, then the SUMOylation site

**Figure 2.6: TDG is destroyed in a DNA damage-, proteasome-, and Cdt2-dependent manner in *Xenopus* egg extract. A)** TDG is destroyed during DNA repair. Damaged plasmid was added to trigger DNA repair and $CRL4^{Cdt2}$-dependent proteolysis. **B)** TDG destruction depends on the proteasome. The proteasome inhibitor MG132 or DMSO (vehicle) was added before the addition of damaged DNA. **C)** Depletion of Cdt2 stabilizes TDG. Cdt2 was immunodepleted from both HSS and NPE before the two extracts were combined and supplemented with TDG and damaged plasmid. For (A) and (B), an extract-specific, non-Cdt1 band detected by the Cdt1 antibody was used as a loading control. For (C), the extract-specific, non-TDG band detected by the TDG antibody was used as a loading control. Benjamin Morris performed the experiment shown in (C).

**Figure 2.6** (continued)



mutant (TDG ΔSUMO) should be destroyed more quickly than its wildtype counterpart. I found that SUMOylation marginally impaired wildtype TDG destruction, as the SUMOylated fraction (Figure 2.7A, B, dotted blue lines) of TDG was destroyed with slower kinetics compared to the unmodified fraction (Figure 2.7A, B, dashed blue lines), but this protective modification was not significant enough to influence destruction of the total TDG pool (SUMOylated+unmodified, Figure 2.7A, B, solid blue line), which was destroyed with near-identical destruction kinetics as TDG ΔSUMO (Figure 2.7A, B, shown in orange). This suggests that SUMOylation is transient and that TDG reassociates with DNA quickly. Destruction of TDG ΔSUMO also depends on damaged DNA, the proteasome, and Cdt2 (Figure 2.8A-B, D).

Because TDG ΔSUMO behaves almost identically to wildtype TDG with respect to destruction kinetics, but it binds to DNA better than the wildtype protein, I used TDG ΔSUMO to study how TDG is modified when it is associated with DNA. To confirm that TDG destruction depends on ubiquitylation when it is bound to DNA, I immobilized MMS-damaged, linear DNA on magnetic beads and added it to egg extract containing methyl-ubiquitin, which prevents polyubiquitin chain formation and destruction but allows monoubiquitylation on multiple sites. Upon recovering the bead-bound DNA, I detected ubiquitylated TDG ΔSUMO (Figure 2.8C). As expected, depletion of Cdt2 eliminated this ubiquitylation (Figure 2.8C).

### 2.3.6 The role of PCNA binding in TDG destruction

The TDG PIP box is not perfect; it is missing one of the two conserved aromatic residues important for productive contact with PCNA. Thus, I next confirmed that TDG destruction depends on its interaction with PCNA. I added a tightly binding PIP box peptide derived from p21 (shown binding PCNA in Chapter 1, Figure 1.2B), another $CRL4^{Cdt2}$ substrate. The PIP box peptide binds to the IDCL of PCNA and decreases RFC-dependent loading of PCNA onto DNA. This peptide also prevents other PIP box-containing proteins from associating with PCNA. As a

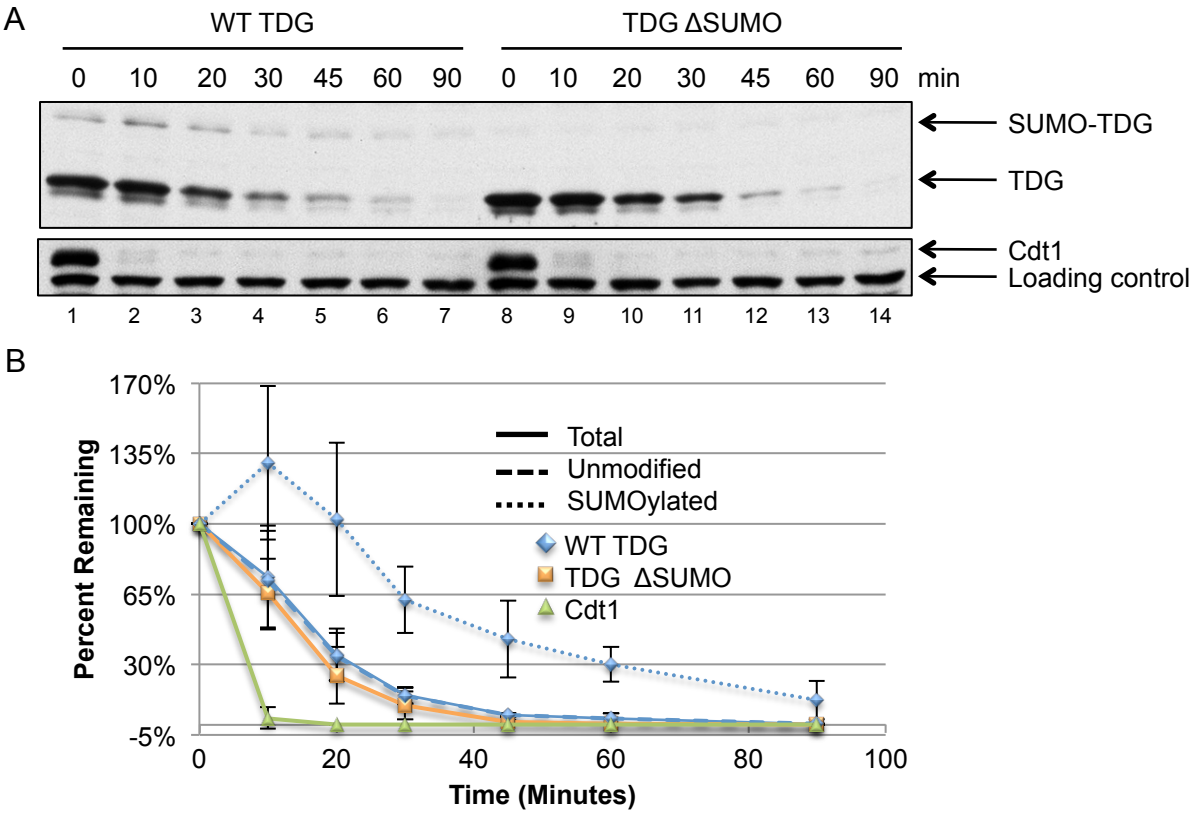
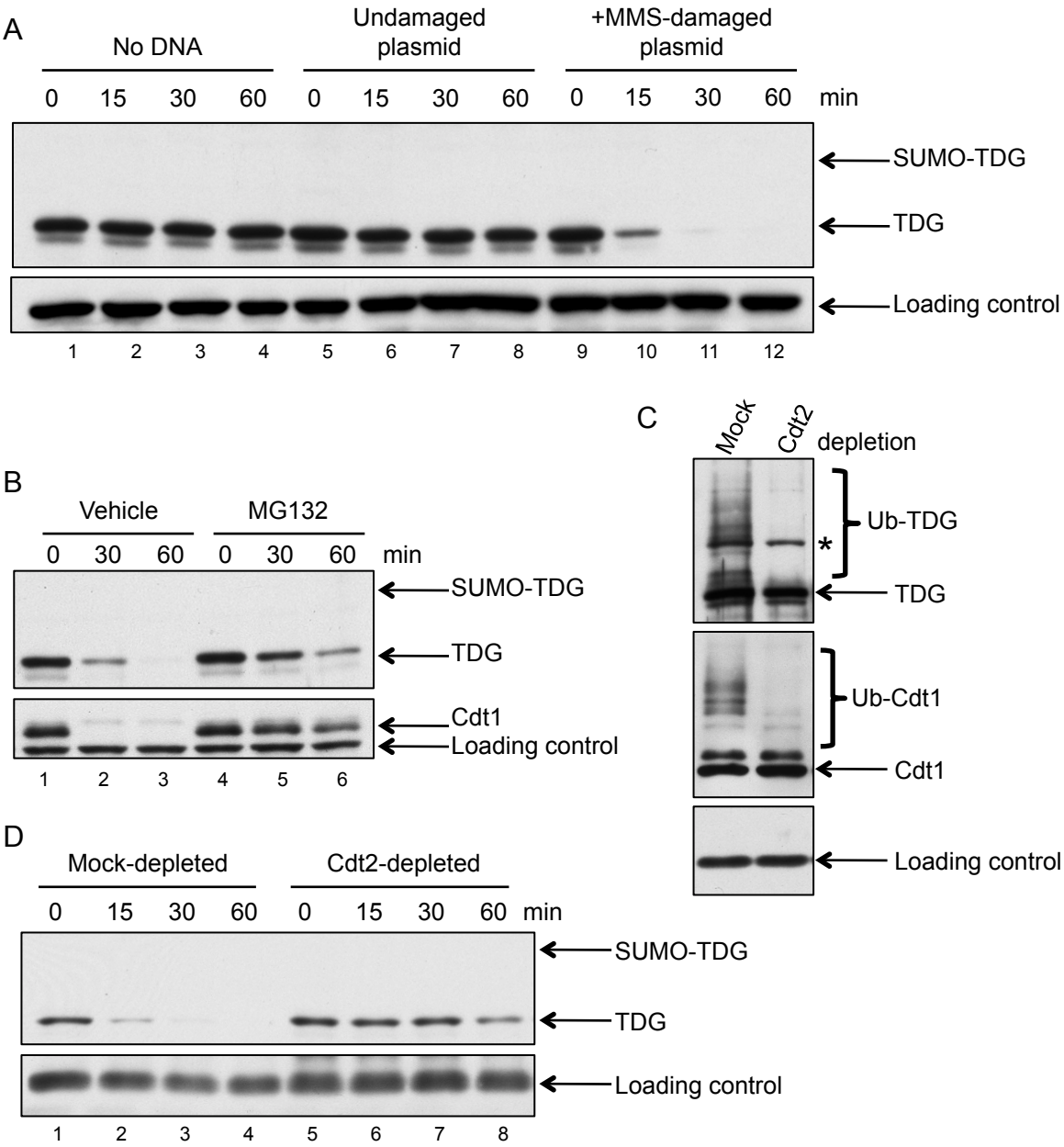



**Figure 2.7: SUMOylation does not affect the destruction kinetics of the total pool of TDG.** **A)** A comparison of wildtype TDG (WT) and TDG ΔSUMO. Recombinant TDG was added to HSS/NPE to a final concentration of 20 nM. Destruction was triggered with 5 ng/μL MMS-damaged plasmid. **B)** Quantification of (A). Quantifications were performed by calculating band density with Image J software. All relevant bands were normalized to a loading control, and the starting concentration was set to 100%. An extract-specific, non-Cdt1 band detected by the Cdt1 antibody was used as a loading control.

**Figure 2.8: TDG ΔSUMO is destroyed in a DNA damage-, proteasome-, and Cdt2-dependent manner in *Xenopus* egg extract. A)** TDG ΔSUMO is destroyed during DNA repair. Damaged plasmid was added to trigger DNA repair and $CRL4^{Cdt2}$-dependent proteolysis. **B)** TDG ΔSUMO destruction depends on the proteasome. The proteasome inhibitor MG132 was added before the addition of damaged DNA. **C)** Depletion of Cdt2 eliminates ubiquitylation of DNA-bound TDG ΔSUMO. Cdt2 was immunodepleted from both HSS and NPE before the two extracts were combined and supplemented with TDG, methyl ubiquitin and linear DNA. The asterisk indicates a band consisting partially of specific signal (ubiquitylated TDG) and partially of non-specific signal. **D)** Depletion of Cdt2 stabilizes TDG ΔSUMO. Extracts were immunodepleted as in (C) and supplemented with TDG and MMS-damaged plasmid. For (A) and (B), an extract-specific, non-Cdt1 band detected by the Cdt1 antibody was used as a loading control. PCNA was used as a loading control for (C). For (D), the extract-specific, non-TDG band detected by the TDG antibody was used as a loading control.

**Figure 2.8** (continued)



control, I used the same peptide with a mutated PIP box. Addition of the PIP box peptide stabilized recombinant TDG, while addition of the control peptide did not (Figure 2.9A). Hence, destruction of wildtype TDG depends on PCNA. The same was true for TDG ΔSUMO (Figure 2.9B). I confirmed that destruction of endogenous TDG also depends on PCNA by monitoring total TDG levels and through TDG immunoprecipitation after the addition of damaged DNA (Figure 2.9C). Endogenous TDG levels dropped upon the addition of damaged DNA (Figure 2.9C, lanes 3-4 for total extract, lanes 9-10 for immunoprecipitated samples), and the PIP box peptide stabilized endogenous TDG (Figure 2.9C, lanes 5-6 for total extract, lanes 11-12 for immunoprecipitated samples).

I next queried the role of TDG's putative PIP degron in damage-dependent destruction. Complete destruction of all known $CRL4^{Cdt2}$ substrates requires a basic residue four amino acids downstream of the PIP box in addition to the PIP box itself (Havens and Walter, 2011). In some substrates, like Set8, other basic residues upstream or downstream of the PIP box also contribute to destruction (Centore et al., 2010). I mutated either the PIP box itself or the downstream basic residues (Figure 2.10A). Importantly, mutation of these downstream residues does not impair the substrate's interaction with the PCNA IDCL (Havens et al., 2012). Mutation of the conserved TDG PIP box (ΔPIP) residues stabilized TDG completely (Figure 2.10, lanes 4-6), confirming that TDG destruction depends entirely on its ability to interact with PCNA. Mutation of the residue four amino acids downstream of the TDG PIP box (ΔB+4) impaired protein destruction, but TDG levels still decreased slightly at later time points, unlike with the PIP box mutant (Figure 2.10, compare lanes 7-9 to lanes 4-6). Mutation of both the B+3 and B+4 residues together (ΔB+3/4) further stabilized TDG (Figure 2.10, lanes 10-12), suggesting that the B+3 residue can function in TDG destruction. These observations support classification of TDG as a $CRL4^{Cdt2}$ substrate.

**Figure 2.9: Blocking the IDCL of PCNA prevents TDG destruction. A)** PCNA inhibition prevents the destruction of wildtype recombinant TDG. **B)** PCNA inhibition prevents the destruction of recombinant TDG ΔSUMO. In all panels, the PIP box peptide was added to a final concentration of 200 μM. For (A) and (B), an extract-specific, non-Cdt1 band detected by the Cdt1 antibody was used as a loading control. **C)** PCNA inhibition prevents the destruction of endogenous TDG. TDG was immunoprecipitated from extract to enrich samples for endogenous TDG, which is present at very low levels.

**Figure 2.9** (continued)

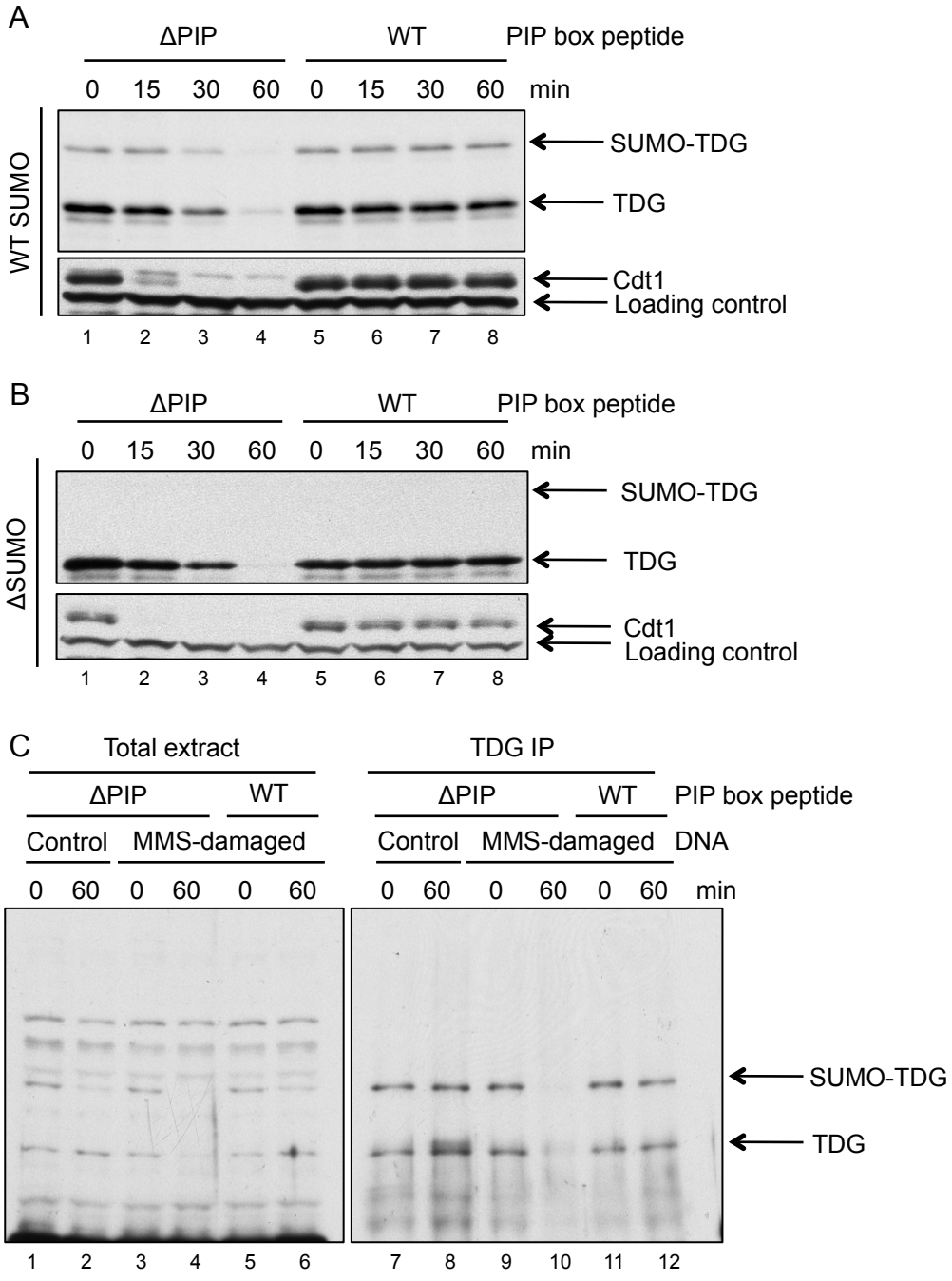
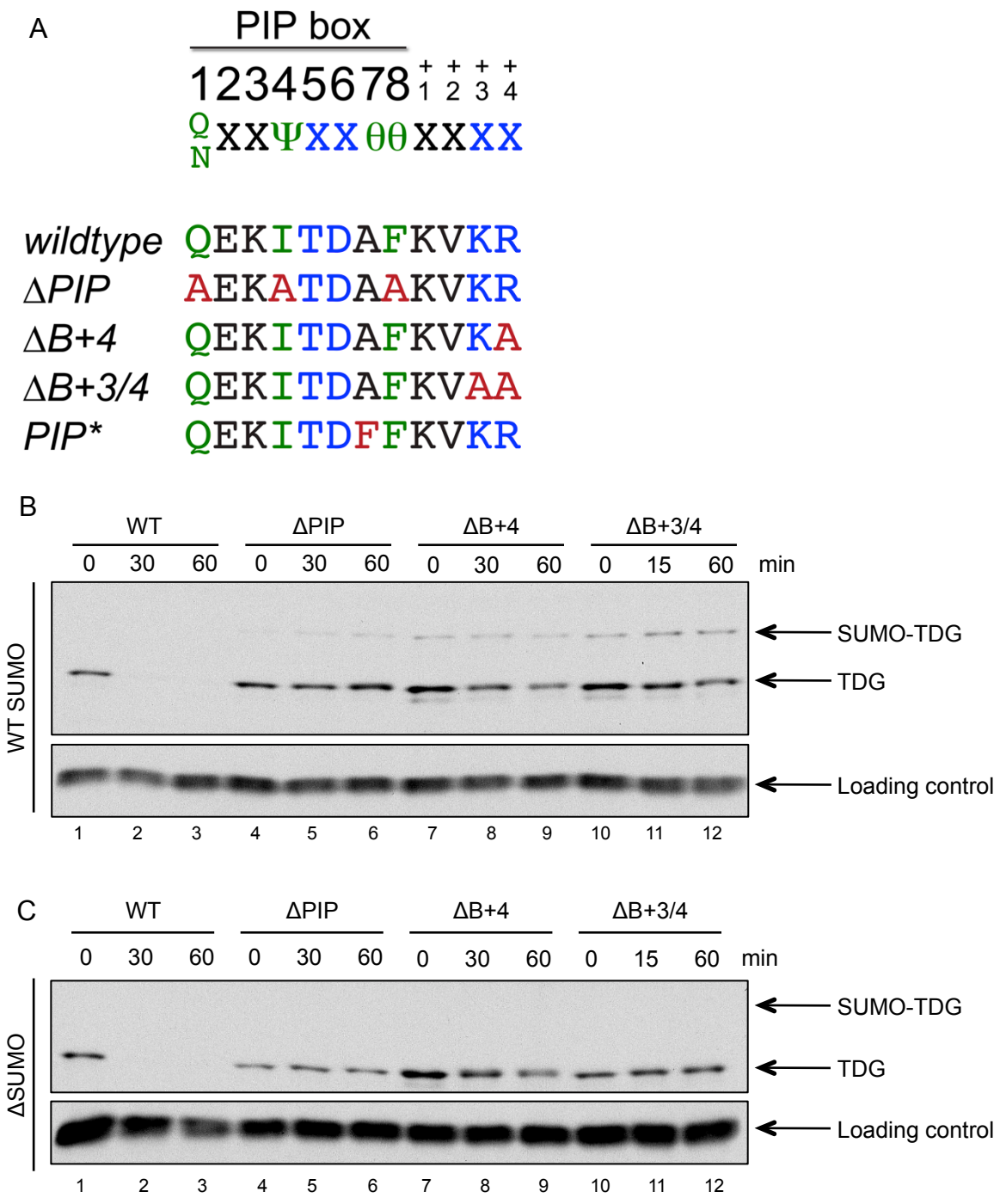


**Figure 2.10: CRL4$^{Cdt2}$-dependent TDG destruction requires the TDG PIP degron. A)** TDG PIP degron mutants (ΔPIP: Q128A/I131A/F135A, ΔB+4: R139A, ΔB+3/4: K138A/R139A, PIP*: A134F) **B)** Mutation of the TDG PIP box or downstream basic residues stabilizes TDG. **C)** Mutation of the same residues in the context of TDG ΔSUMO leads to stabilization. Note that mutation of the B+4 residue alone does not cause complete stabilization. The extract-specific, non-TDG band detected by the TDG antibody was used as a loading control.

**Figure 2.10** (continued)

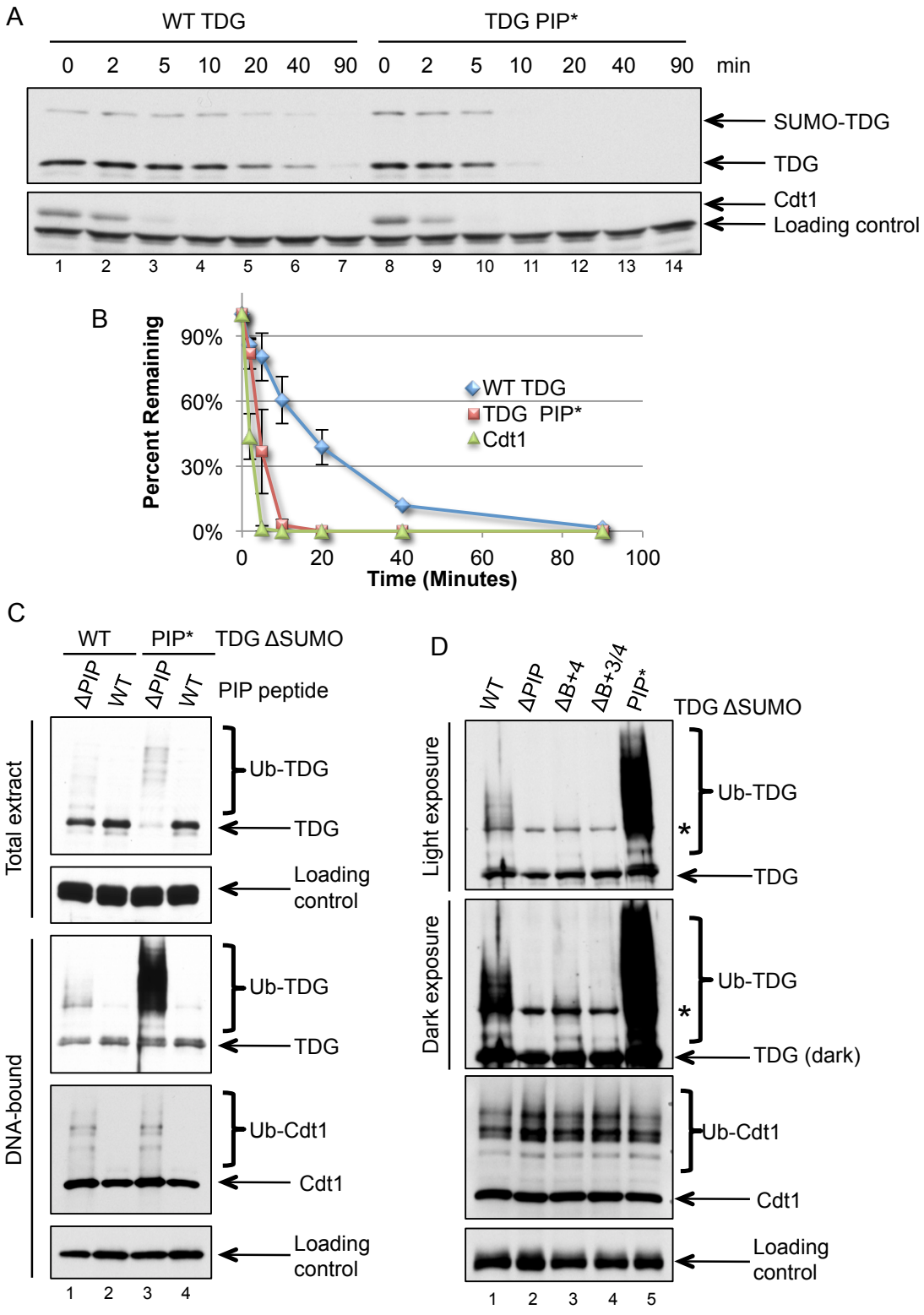


I postulated that the slower destruction kinetics of TDG compared to Cdt1 could be due to its imperfect PIP box. TDG lacks one of the typically conserved aromatic residues in its PIP box, so I mutated the alanine residue at the 7 position of TDG's PIP box to phenylalanine (A134F) to improve the PIP degron (PIP*, Figure 2.10A). TDG PIP* was destroyed much faster than wildtype TDG, with kinetics much closer to those of Cdt1 (Figure 2.11). Thus, an imperfect PIP box can still trigger $CRL4^{Cdt2}$-dependent destruction, but with slower kinetics than those that associate with PCNA better than wildtype TDG. Improving TDG's PIP box allowed it to access PCNA more efficiently despite the presence of other PCNA-binding proteins, in effect leading to faster destruction. Hence, $CRL4^{Cdt2}$ prioritizes substrates based on their ability to associate with PCNA. In *Xenopus* egg extract, $CRL4^{Cdt2}$ prefers substrates with a perfect PIP degron like Cdt1, but it is still capable of targeting lesser substrates during DNA repair.

Moreover, improving the TDG PIP box led to greater ubiquitylation on DNA (Figure 2.11C, lane 3). Addition of the short PIP box-containing peptide of p21 eliminated this ubiquitylation (Figure 2.11C, lane 4) and that of the wildtype protein (Figure 2.11C, lane 2), indicating the ubiquitylation is most likely PCNA-dependent. Mutation of the conserved PIP residues eliminated all TDG ubiquitylation (Figure 2.11D, lane 2), and mutation of the B+4 residue led to decreased ubiquitylation (Figure 2.11D, lane 3). Additional mutation of the B+3 residue eliminated TDG ubiquitylation (Figure 2.11D, lane 4). I confirmed that mutation of the degron residues also stabilized TDG ΔSUMO and that improvement of the PIP box triggered its more rapid destruction (Figure 2.10C and data not shown). Therefore, TDG is a $CRL4^{Cdt2}$ substrate during DNA damage in both human cells and *Xenopus* egg extract, but $CRL4^{Cdt2}$ preferentially targets substrates with a perfect PIP box.

**Figure 2.11: Improving TDG's PIP box results in faster destruction and increased ubiquitylation on DNA. A)** MMS-damaged plasmid was added to a final concentration of 5 ng/μL in HSS/NPE and destruction of WT TDG and TDG PIP* was monitored. **B)** Quantification of (A). TDG destruction was calculated as in Figure 2.7. Improvement of the TDG PIP box led to more rapid destruction, similar to Cdt1 destruction. **C)** Improving the TDG PIP box leads to more robust ubiquitylation of TDG on chromatin. Addition of a PIP box competitor peptide abolished ubiquitylation of both wildtype TDG and TDG PIP*. **D)** Mutation of TDG's conserved PIP degron residues eliminates TDG ubiquitylation on DNA. Conversely, improving the PIP box improves ubiquitylation. The extract-specific, non-Cdt1 band detected by the Cdt1 antibody was used as a loading control for (A). Orc2 was used as a loading control for (C), and RPA-14 was used as a loading control for (D).

**Figure 2.11** (continued)



### 2.3.7 TDG destruction during S phase

$CRL4^{Cdt2}$ substrates are also destroyed during the transition into S phase. Indeed, TDG is destroyed during S phase in human cells (Hardeland et al., 2007) and was stabilized in response to siCdt2 (Figure 2.3). I therefore monitored TDG levels during DNA replication in egg extracts. Unlike Cdt1, wildtype TDG and TDG ΔSUMO remained stable during the course of DNA replication in HSS/NPE (Figure 2.12A, lanes 1-4 and D, lanes 1-4). In contrast, TDG PIP* and TDG ΔSUMO/PIP* were destroyed rapidly in a replication-dependent manner during DNA replication (Figure 2.12A, lanes 5-12 and D, lanes 5-12), showing that TDG can serve as a $CRL4^{Cdt2}$ substrate during DNA replication in *Xenopus* egg extract when it can access PCNA through an improved PIP box.

Though total wildtype TDG levels did not change during DNA replication in extract, it is possible that TDG is destroyed locally on DNA when the replication fork approaches, and failure to destroy TDG could generate a replication fork block because TDG binds so tightly to abasic sites (its catalytic product). To study the effects of TDG on DNA replication, I monitored replication of a plasmid containing a G-T mispair, which is the cognate substrate for TDG, in the presence and absence of high levels of TDG. To eliminate possible $CRL4^{Cdt2}$-independent reduction of DNA-bound TDG, I performed these experiments using TDG ΔSUMO. TDG acted on this G-T mispaired plasmid substrate in a base release assay in either activity buffer or egg lysis buffer, which is most similar to extract buffer conditions (Figure 2.5A, lanes 6-7). However, TDG ΔSUMO/ΔPIP had no effect on bulk DNA replication of the plasmid containing a G-T mispair or on the resolution of replication intermediates, despite the fact that it should not be removed from DNA via SUMOylation or via $CRL4^{Cdt2}$-dependent destruction (Figure 2.13A, B). I also looked for transient fork stalling at the site of the mispair and found no pausing above background levels (Figure 2.13C, D). I therefore questioned whether TDG acts on the mispair in *Xenopus* egg extract. After one hour in extract, the mispair remained unrepaired (Figure 2.13E),

**Figure 2.12: Replication-dependent TDG destruction requires an improved PIP box in *Xenopus* egg extract. A)** TDG was added to extract at the start of replication (when NPE was added to HSS containing licensed plasmid), and TDG levels were monitored during replication. Wildtype TDG remained stable, while TDG PIP* was destroyed. Geminin was added to HSS to inhibit licensing. **B)** Corresponding DNA replication during the destruction assay shown in (A). **C)** Quantification of (B). **D-F)** The same experiment was repeated with TDG ΔSUMO. Note that the slight increase in DNA replication in (C) "PIP*+geminin, 10 minutes" is due to signal from the ΔPIP 60 minute lane and not due to actual replication. The extract-specific, non-Cdt1 band detected by the Cdt1 antibody was used as a loading control. In (B) and (E), RI stands for replication intermediates, N stands for nicked, and S stands for supercoiled.

**Figure 2.12** (continued)

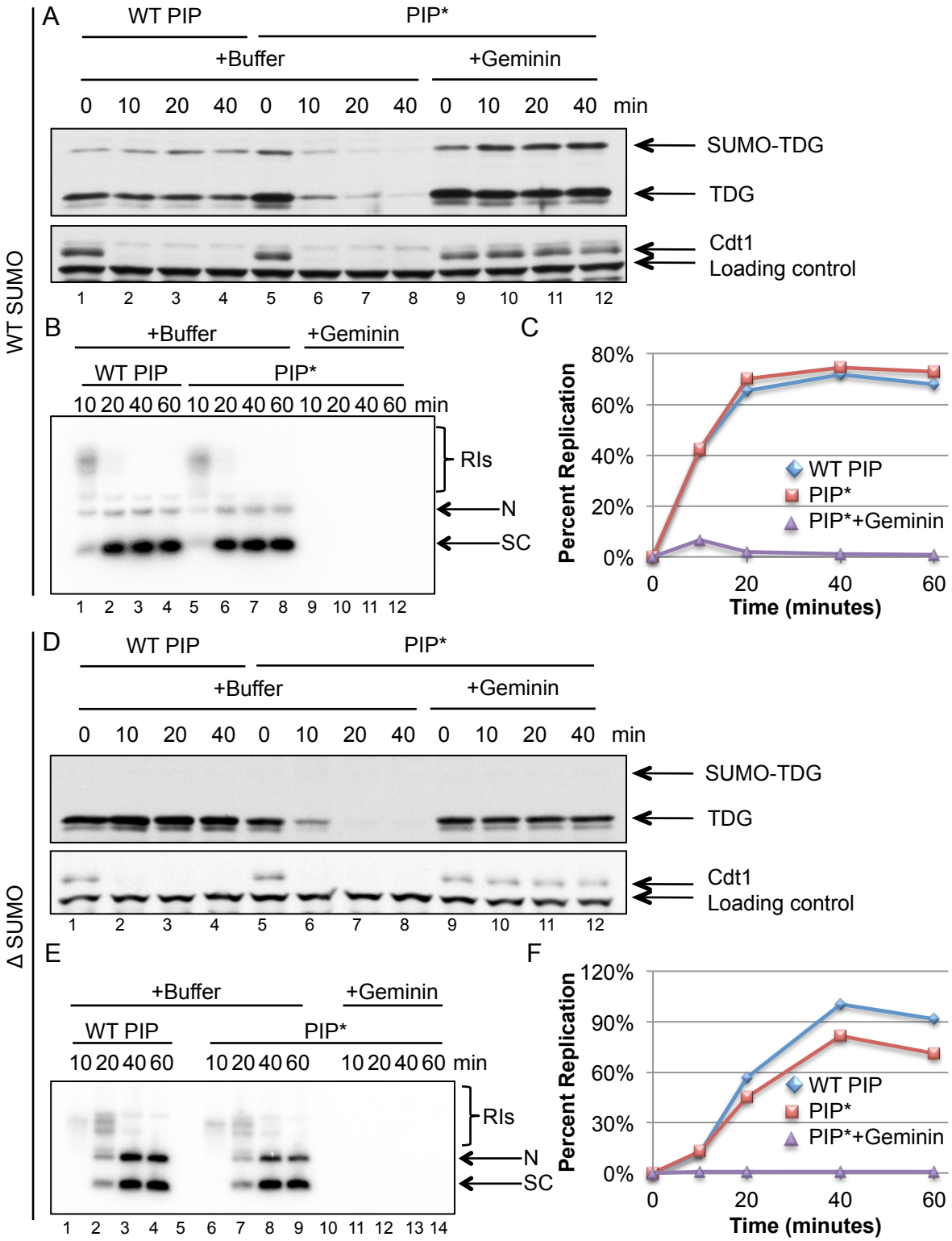
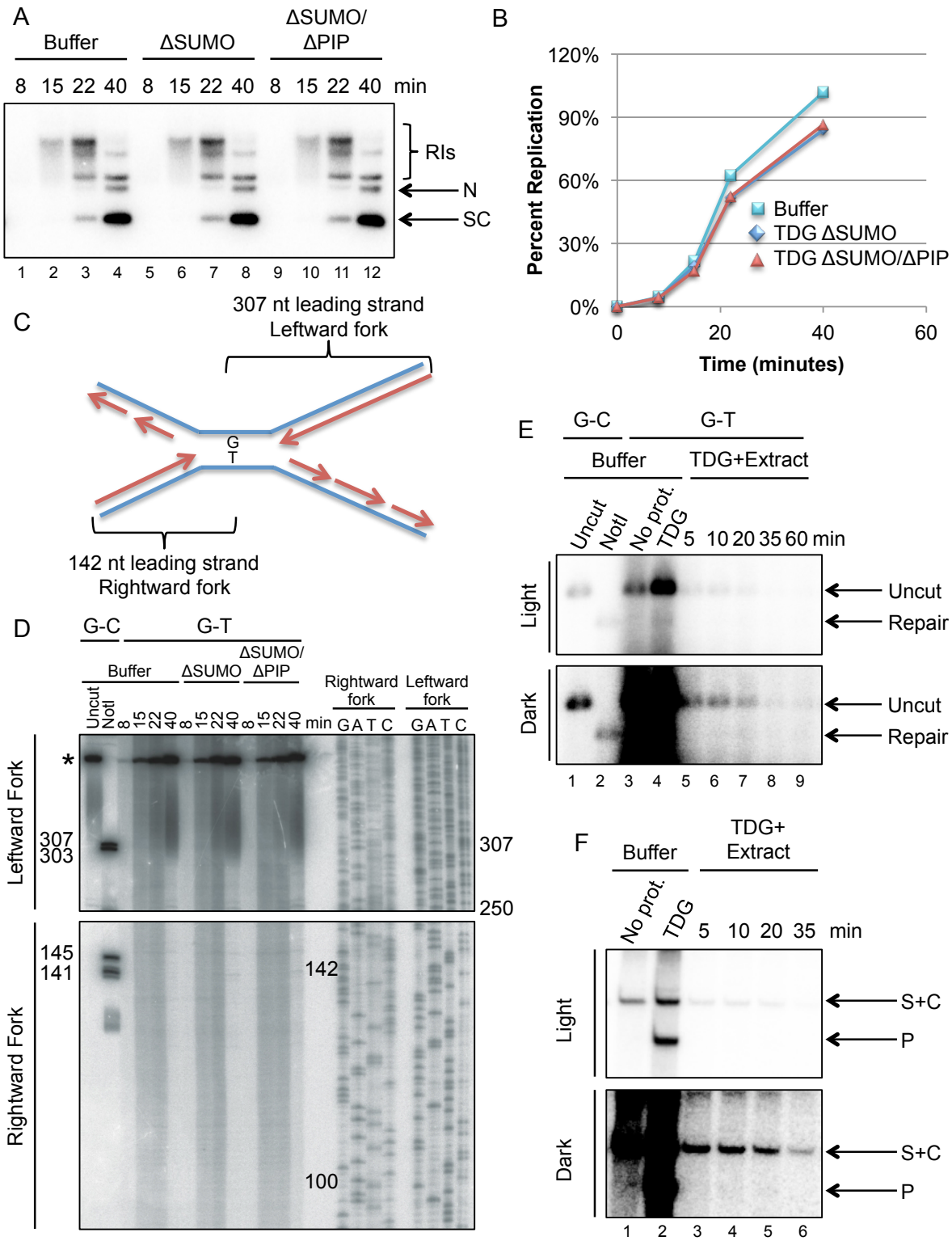


**Figure 2.13: High levels of TDG ΔSUMO/ΔPIP do not affect DNA replication in *Xenopus* egg extract. A)** 250 nM TDG or TDG ΔSUMO/ΔPIP was added to HSS/NPE containing a licensed plasmid containing a G-T mispair (see Figure 2.1), and bulk replication was monitored. **B)** Quantification of (A). **C)** Approach used to monitor for transient stalling triggered by TDG at the site of the G-T mispair. The mispaired plasmid was digested with PvuII, as in Figure 2.1. **D)** Transient fork stalling was monitored with a sequencing gel that measures accumulation of the nascent strand at various lengths. TDG did not affect fork progression.

**Figure 2.13** (continued)



lanes 5-9) and did not generate an abasic site (Figure 2.13F, lanes 3-6) in extract despite the extremely high TDG concentration that triggers base hydrolysis in the standard activity buffer and ELB (Figure 2.5A, lanes 6-7). Experiments addressing TDG activity must be expanded and repeated to be conclusive, but this preliminary data suggests that a redundant mechanism might restrict TDG access to or activity on DNA during replication in *Xenopus* egg extract, even when TDG is presented with its cognate DNA substrate. In summary, I have found that failure of $CRL4^{Cdt2}$ to trigger destruction of TDG does not lead to replication defects, but I cannot rule out a role for $CRL4^{Cdt2}$ in preventing TDG-dependent problems in the absence of unidentified forms of TDG regulation.

### 2.3.8 Checkpoint signaling is not required for TDG destruction

In egg extract, *Xenopus* TDG requires an improved PIP box for S phase-specific proteolysis, while in tissue culture, human TDG is destroyed despite its imperfect degron. I reasoned that checkpoint signaling, one major difference between the nuclear environments of *Xenopus* egg extract and human cells, might be necessary for TDG destruction. In response to damaged DNA, *Xenopus* egg extract mounts a checkpoint response that is not present during DNA replication (Clute and Masui, 1997; Conn et al., 2004; Dasso and Newport, 1990; Hensey and Gautier, 1998; Kappas et al., 2000; Kimelman et al., 1987; Kumagai and Dunphy, 1999; Kumagai et al., 1998). Thus, I inhibited checkpoint signaling with caffeine during DNA repair and monitored TDG levels. Damage-dependent TDG ΔSUMO destruction did not change upon the addition of caffeine, despite diminished ATR checkpoint signaling as measured by Chk1 phosphorylation (Figure 2.14, compare lanes 5-8 to lanes 1-4). TDG ΔSUMO/ΔPIP was stable regardless of the presence of caffeine (Figure 2.14, lanes 9-14). Therefore, TDG destruction does not require checkpoint signaling, and differences in checkpoint signaling between S phase human cells and replicating *Xenopus* egg extract do not explain the $CRL4^{Cdt2}$ requirement for an improved TDG PIP box in *Xenopus* egg extract.

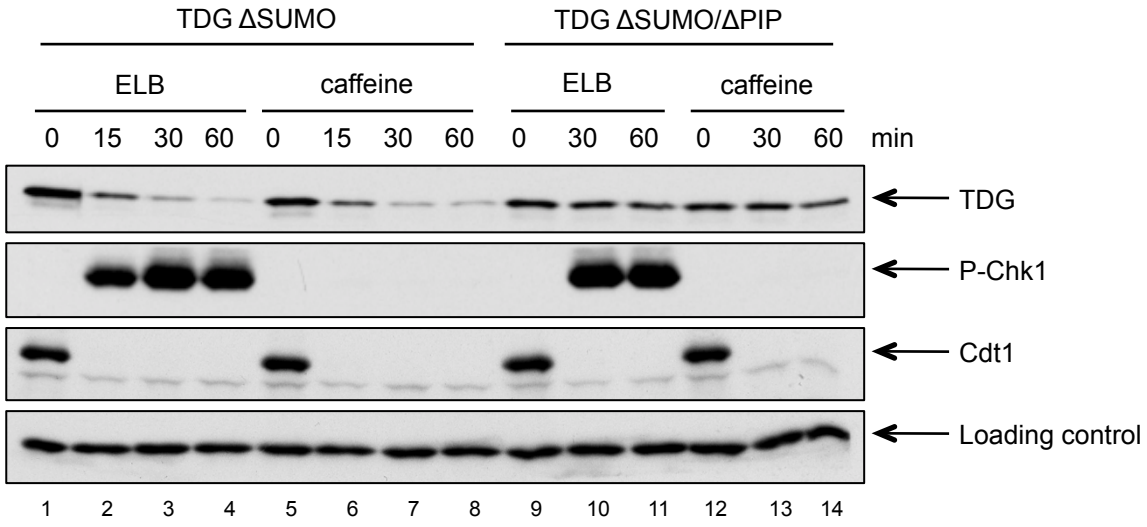



**Figure 2.14: TDG proteolysis does not require checkpoint signaling.** 3 mM caffeine was added to inhibit ATR, ATM, and DNA-$PK_{CS}$ activity during DNA repair. TDG destruction occurred normally in the presence of caffeine. The extract-specific, non-Cdt1 band detected by the Cdt1 antibody was used as a loading control.

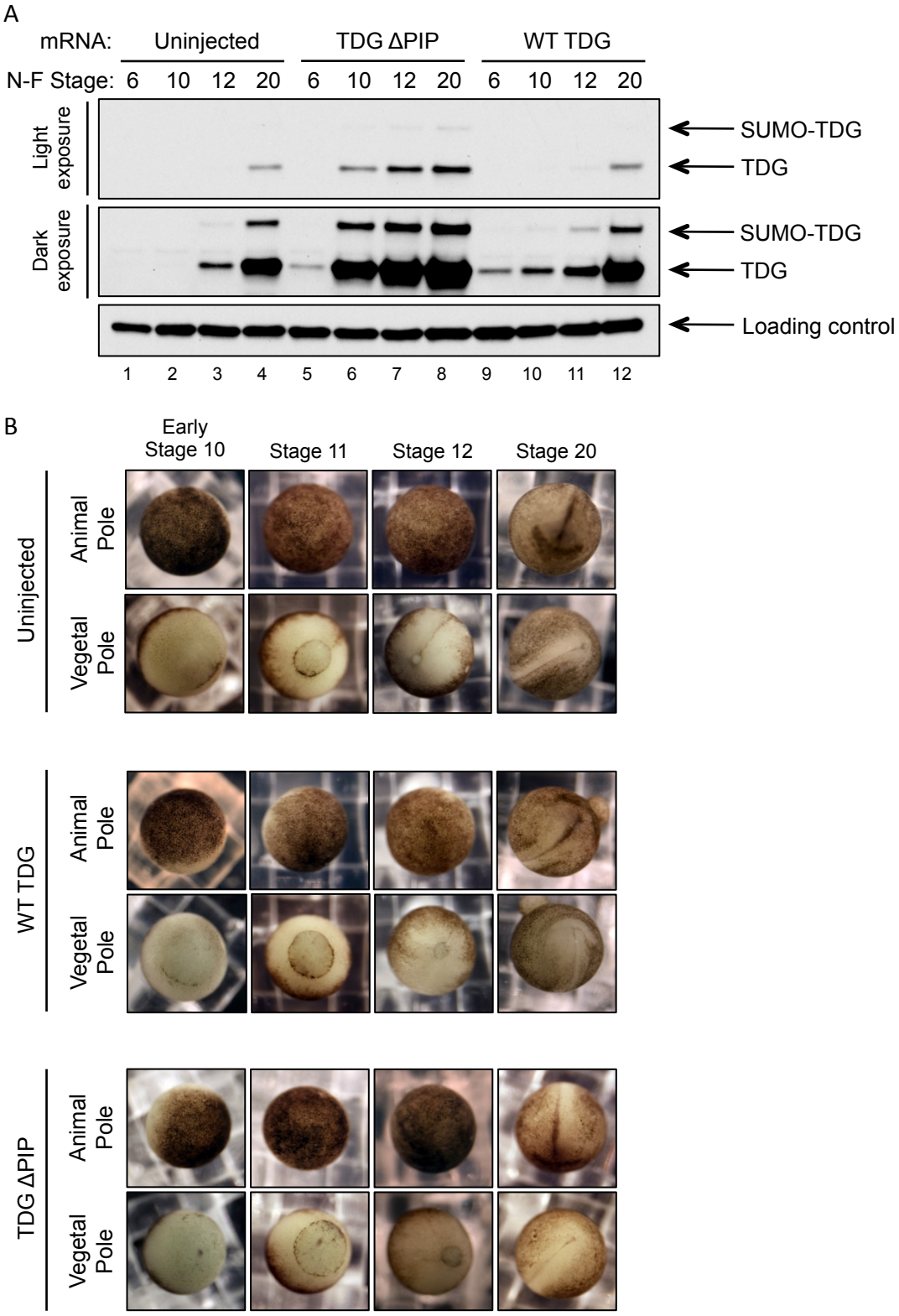
### 2.3.9 The physiological relevance of $CRL4^{Cdt2}$-dependent TDG destruction

Another possible explanation for the differences in S phase-specific proteolysis between *Xenopus* egg extract and human cells is that $CRL4^{Cdt2}$-dependent TDG proteolysis is developmentally regulated. Since TDG is maintained at very low levels until mid-gastrulation in developing embryos (Figure 2.4E) we reasoned that $CRL4^{Cdt2}$ might not destroy TDG until later in development, when TDG is normally present. Only after $CRL4^{Cdt2}$-dependent TDG proteolysis becomes active would we expect the degron mutant to cause problems that wildtype TDG does not. In mice, the essential role of TDG is related to its function in DNA demethylation. Mice lacking TDG suffer from aberrant DNA methylation and repression of genes involved in tissue differentiation (Cortázar et al., 2011; Cortellino et al., 2011). Moreover, forced TDG recruitment to specific genes leads to localized demethylation (Gregory et al., 2012). I therefore rationalized that failure to destroy TDG during S phase could cause aberrant demethylation of certain genes, which in turn might induce epigenetic changes and affect embryonic development. To study this possibility, I microinjected each cell of Stage 2 *Xenopus* embryos with either wildtype TDG or TDG ΔPIP mRNA and monitored them at different stages of development.

As expected, embryos injected with wildtype TDG mRNA contained more TDG than uninjected controls, and protein levels steadily increased throughout early development (Figure 2.15A, compare lanes 1-4 to lanes 9-12). Levels of wildtype TDG and TDG ΔPIP protein were similar during the early cleavage divisions (wildtype and ΔPIP levels are identical at Stage 6 in Figure 2.15A, compare lanes 5 and 9). However, after the blastula stage, TDG ΔPIP accumulated to much higher levels than wildtype TDG (Figure 2.15A, compare lanes 6-8 to lanes 10-12). Given that the microinjected RNA is identical except at the mutated nucleotides encoding PIP box residues and that a PIP box has not been shown to affect translation, it is unlikely that this

**Figure 2.15: TDG ΔPIP is stabilized in *X. laevis* embryos after the early cleavage divisions but does not affect development. A)** Embryos microinjected with TDG mRNA were harvested and lysed at the stages indicated. Embryos injected with wildtype TDG mRNA express similar levels of TDG protein as embryos injected with an equivalent amount of TDG ΔPIP at N-F Stage 6. TDG ΔPIP accumulates to much higher levels than WT TDG by early gastrulation (N-F Stage 10). WT TDG-injected embryos express higher TDG levels than uninjected controls. PCNA was used as a loading control. **B)** Embryos were microinjected as in (A) and monitored for developmental defects. Embryos expressing TDG ΔPIP developed normally into the tadpole stages despite accumulation of degron-mutant TDG.

**Figure 2.15** (continued)



accumulation is due to differences in RNA translation. Instead, this accumulation is consistent with a failure to destroy TDG ΔPIP at times when wildtype TDG is destroyed, suggesting that TDG becomes a $CRL4^{Cdt2}$ substrate during normal cell cycles in developing *Xenopus* embryos some time between the blastula and early gastrula stages. Benjamin Morris of the Walter lab plans to confirm that injected mRNA levels were equal between WT TDG and TDG ΔPIP through qPCR in the future.

I closely followed the development of microinjected embryos from fertilization into the tadpole stage. However, I observed no developmental defects, despite a large excess of TDG ΔPIP over TDG WT, indicative of a failure to destroy TDG during S phase (Figure 2.15B, showing the animal and vegetal poles of single embryos from stages 10-20). Therefore, $CRL4^{Cdt2}$-dependent TDG proteolysis is not required for normal *X. laevis* development through the early tadpole stage. It is possible that TDG proteolysis is required for later development, but injected mRNAs only persist in embryos for approximately two days (Blitz et al., 2006). Although the TDG ΔPIP protein is stable during S phase, it will not persist indefinitely, as proteins are ultimately turned over regardless of cell cycle-specific proteolysis. Therefore, we cannot use this approach to study the effects of stabilizing TDG during later developmental stages.

I have now identified two new forms of TDG regulation: 1) TDG abundance is regulated at the translational or post-translational level early in development, since mRNA levels are high even in the unfertilized egg, but protein levels are very low; and 2) TDG is a $CRL4^{Cdt2}$ substrate during DNA repair in egg extract and during normal cell cycles of *X. laevis* embryos after the blastula stage. Additionally, previous reports have shown that SUMOylation, acetylation, and phosphorylation control the access of TDG to DNA (Baba et al., 2005; Hardeland et al., 2002; Mohan et al., 2010; Smet-Nocca et al., 2011). With so many different types of regulation to restrict TDG's access to DNA, I reasoned that TDG must be detrimental to DNA replication or

negatively affect another S phase-specific process. However, I could not identify any replication or developmental defects upon stabilizing TDG.

## *2.4 Discussion*

In recent years, TDG has emerged as a key player not only in DNA repair but also in regulating epigenetic inheritance through demethylation. TDG is an essential gene, likely due to its role in regulating DNA methylation, and several theories have emerged as to how TDG contributes to establishing and maintaining methylation patterns on DNA. Several forms of TDG regulation have been discovered: 1) SUMOylation facilitates TDG turnover on DNA, 2) both acetylation and phosphorylation regulate its access to DNA, and 3) S phase-specific proteolysis eliminates TDG during DNA replication. I have identified new mechanisms of TDG regulation. I found that TDG is destroyed in a DNA damage-, proteasome-, Cdt2-, PCNA-, and PIP-degron-dependent manner in *Xenopus* egg extract, making it a *bona fide* $CRL4^{Cdt2}$ substrate. Regulation by $CRL4^{Cdt2}$, which targets its substrates for destruction during both DNA repair and DNA replication, explains S phase-specific TDG proteolysis that has been previously reported. However, during the rapid cell cycles during early cleavage divisions and in *Xenopus* egg extract, TDG required an improved degron to be destroyed in S phase. Nonetheless, a wildtype PIP degron was sufficient to prevent aberrant accumulation of TDG after the early cleavage divisions during *Xenopus* development, suggesting that $CRL4^{Cdt2}$ begins to target TDG during normal cell cycles in the blastula or early gastrula stages. Before the early gastrula stages, TDG levels are also regulated through another undefined pathway that limits protein expression despite relatively high mRNA levels.

Somewhat surprisingly, I found no detrimental effects of stabilizing TDG on DNA replication or development in *Xenopus*. One possibility is that another redundant mechanism regulates TDG, so the effect of stabilizing TDG was masked in our experiments. In the case of the well-

characterized $CRL4^{Cdt2}$ substrate Cdt1, multiple redundant pathways regulate its activity after initiation of DNA replication. Defects in replication (re-replication, in the case of Cdt1) only occur when more than one regulatory pathway is compromised. It is therefore possible that unknown redundant pathways prevent TDG from acting on DNA during replication, and TDG only causes defects during S phase if TDG overcomes other regulators.

All known substrates of $CRL4^{Cdt2}$ are in some way toxic during S phase (Havens and Walter, 2011), but failure to destroy TDG had no effect on progression through replication. However, recombinant TDG is not active in extract on its cognate substrate, a G-T mispair, at a concentration where it is active in multiple buffer conditions, suggesting that another unknown form of regulation might restrict it from DNA or block its activity in egg extract. It is possible that eliminating $CRL4^{Cdt2}$-dependent regulation in addition to other unknown forms of regulation would affect DNA replication. At this point, I can only speculate as to why TDG is destroyed during S phase. High overexpression of TDG (20-30-fold) causes S phase arrest in 293T cells, and such high overexpression blocks cell proliferation (Hardeland et al., 2007). It is possible that product-inhibited TDG, which binds tightly to an abasic site, causes replication blocks that cannot be resolved in S phase. We have several theories as to why TDG destruction would be evolutionarily favorable. For example, TDG cannot distinguish the parental strand from the nascent strand during replication, so its destruction would allow the mismatch repair system sole access to G-T mispairs that arise due to replication error. Mismatch repair is better equipped to handle mistakes during replication because it can discern the newly synthesized DNA strand. If we can uncover conditions that support TDG activity in egg extract, we can easily test this theory by monitoring the effect of TDG on the error rate during replication, especially in CpGs.

Alternatively, the presence of TDG on DNA during replication could disrupt epigenetic

inheritance by interfering with proper DNA methylation or demethylation in S phase. In support of this idea, the putative PIP degron sequence is highly conserved throughout higher eukaryotes that regulate epigenetic inheritance through DNA methylation. It is absent from organisms that display low or no DNA methylation, like *Drosophila* (Figure 2.2 and Cortázar et al., 2007). TDG homologues in lower eukaryotes cannot process thymine or 5-methylcytosine derivatives and do not contain a PIP box (Hardeland et al., 2003). In mice, the essential role of TDG is likely related to its function in epigenetic regulation (Cortázar et al., 2011; Cortellino et al., 2011), so its destruction could be necessary to grant another pathway control over proper inheritance of DNA methylation during S phase without interference from TDG. Again, if we can develop conditions that support TDG activity in egg extract, we can test this theory by adding methylated DNA and monitoring its methylation state and the methylation state of the nascent strand during DNA replication.

TDG acts on mispairs that arise from modification of cytosine and 5meC by other proteins, including activation-induced deaminase (AID) and Tet proteins. Cells also contain another G-T specific glycosylase, MBD4, that associates with methylated DNA in heterochromatin regions (Sjolund et al., 2012). During DNA replication, heterochromatin must decondense to allow passage of replication forks. It is possible that during DNA replication, TDG could gain aberrant access to these regions normally maintained by MBD4, triggering aberrant demethylation and activation of silent genes. $CRL4^{Cdt2}$ might be one of multiple redundant mechanisms that restrict TDG access to these regions.

Finally, regardless of the identification of defects due to TDG stabilization, the knowledge that TDG is a $CRL4^{Cdt2}$ substrate can be applied to the medical field. TDG is responsible removal of 5-fluorouracil (5-FU) from DNA by BER (Cortázar et al., 2007; Hardeland et al., 2000; 2003; Kunz et al., 2009). 5-FU is a widely used chemotherapeutic drug; it is one of the primary

treatments for colorectal and breast cancers (Longley et al., 2003). Upon entry into cells, 5-FU is converted to a number of different metabolites that cause S phase arrest and could potentially trigger cancer cell death in several ways. 5-FU inhibits thymidylate synthase, in effect causing decreased thymidine levels and increased uracil levels (Longley et al., 2003). This not only impairs DNA replication but also leads to frequent misincorporation of both uracil and 5-FU into DNA. Because our understanding of 5-FU action is vague and 5-FU is so frequently used in cancer treatment, understanding what factors contribute to 5-FU resistance and sensitivity is important to improve cancer treatment in the future.

A recent study found that TDG is important for 5-FU sensitivity in mouse embryonic fibroblasts and HeLa cells (Kunz et al., 2009). Cells lacking TDG had much higher survival rates upon treatment with the drug, and cells expressing additional TDG showed greater sensitivity (Kunz et al., 2009). It is possible that by acting on so many 5-FU lesions, TDG and the BER pathway cause an irreparable number of DNA nicks that lead to double-stranded breaks when they are in close proximity to each other. Though these experiments must be repeated in the cell lines of cancers typically treated with 5-FU to confirm medical relevance, the link between TDG and 5-FU sensitivity is promising. If the relationship is confirmed, then stabilizing TDG in patients undergoing 5-FU treatment could allow for lower doses of the drug and decreased resistance. The new knowledge that TDG is a $CRL4^{Cdt2}$ substrate adds the possibility of modifying an existing cancer drug regimen to maximize 5-FU sensitivity. A general CRL inhibitor, MLN4924 (Millennium Pharmaceuticals), is currently in clinical trials as a chemotherapeutic agent (information publicly available from the National Cancer Institute). MLN4924 inhibits the activation of all cullin ring ligases, in effect preventing proteolysis during many cellular processes (Soucy et al., 2009). MLN4924 would therefore be expected to stabilize TDG through inhibition of $CRL4^{Cdt2}$. Therefore, combination therapy with 5-FU could be a successful cancer treatment regimen. This possibility could easily be tested in human cells.

# Chapter 3

# How the homotrimeric nature of PCNA contributes to $CRL4^{Cdt2}$ function

## *3.1 Introduction*

Unlike other ubiquitin ligases, $CRL4^{Cdt2}$ functions specifically during DNA replication and repair to downregulate many proteins whose presence in S phase is detrimental. The best-characterized substrate of $CRL4^{Cdt2}$ is Cdt1 (Jin et al., 2006), which functions to recruit the replicative MCM2-7 helicase during origin licensing. After cells progress into S phase, Cdt1 activity could facilitate detrimental origin licensing on regions of the genome that have already been replicated (Arias and Walter, 2005). For this reason, $CRL4^{Cdt2}$ marks Cdt1 for destruction upon entry into S phase. $CRL4^{Cdt2}$ also targets Cdt1 for destruction in response to DNA damage, though the reasons are unclear (Hu et al., 2004).

$CRL4^{Cdt2}$ substrates must bind PCNA through their PIP boxes (Figure 1.2) to become degradation targets (Arias and Walter, 2006), but $CRL4^{Cdt2}$ does not target all PIP box-containing proteins for destruction; in fact, most PCNA-interacting proteins are stable when bound to PCNA. All known $CRL4^{Cdt2}$ substrates contain a basic residue four amino acids downstream of the PIP box (the B+4 position relative to the PIP box). This sequence contributes to a “PIP degron.” (Havens and Walter, 2009). PCNA-interacting proteins whose stability is not regulated by $CRL4^{Cdt2}$ do not contain a complete PIP degron (Chapter 1, Figure 1.2). The binding surface $CRL4^{Cdt2}$ includes residues on both the substrate and on DNA-bound PCNA (Chapter 1, Figure 1.2 and Havens and Walter, 2009).

$CRL4^{Cdt2}$ detects its substrates only when they are bound to PCNA loaded on DNA. Residues on both PCNA and on the substrate itself recruit Cdt2, the putative substrate specificity factor for the ubiquitin ligase (Havens and Walter, 2009; Havens et al., 2012). Our current model envisions that a substrate binds one monomer of DNA-bound PCNA

via its PIP box, and the substrate-PCNA complex then recruits $CRL4^{Cdt2}$ (Chapter 1, Figure 1.3A). Interestingly, a recent report from the Yew lab suggests that Cdt2 can bind PCNA directly, independently of its substrate (Kim et al., 2010). It is therefore possible that within the PCNA homotrimer, Cdt2 contacts two subunits: one via direct contacts with the substrate-IDCL complex and another through direct contacts between Cdt2 and an IDCL of an adjacent PCNA subunit (Chapter 1, Figure 1.3B). This second model predicts that $CRL4^{Cdt2}$ can only function with PCNA that has at least two wildtype subunits with intact IDCLs.

Because of the observation from the Yew lab that Cdt2 binds PCNA independently of its substrate (Kim et al., 2010), I am particularly interested in determining whether the "tool belt" model applies during $CRL4^{Cdt2}$ function. The tool belt model for PCNA function—in which PCNA binds three separate proteins through its three separate subunits to allow for rapid exchange of DNA replication and repair proteins as they are needed on DNA—is widely referenced and accepted in the field (Freudenthal et al., 2011; Indiani et al., 2005; Sutton, 2010; Zhuang and Ai, 2010). However, there has been no way to distinguish this model from other models, like the "dynamic hand-off model" (Prosperi, 2006) because replacement of individual PCNA subunits within the homotrimeric ring has not been possible.

Here, I address how many functional PCNA subunits are required for $CRL4^{Cdt2}$ function and DNA replication. I generated a single chain polypeptide in which all three PCNA subunits are connected through flexible linker sequences ("single chain PCNA," or $PCNA^{SC}$), and determined how multiple IDCLs on PCNA contribute to $CRL4^{Cdt2}$ function. I found that a single IDCL was sufficient for modest destruction of the well-characterized $CRL4^{Cdt2}$ substrate Cdt1, but two separate IDCLs were necessary to support complete Cdt1 destruction. I was unable to determine whether this observation is due to

decreased substrate binding sites on $PCNA^{SC}$ with only one wildtype IDCL or due to a requirement for multiple functional subunits on the same PCNA trimer for $CRL4^{Cdt2}$ function. Additionally, I found that a single subunit is sufficient for leading strand elongation, suggesting that the tool belt model does not apply for this process. Implications and future use of this new molecular tool, $PCNA^{SC}$, are discussed.

## *3.2 Materials and Methods*

Sperm chromatin, HSS, and MMS-damaged DNA were prepared as described above (Chapter 2). Damaged linear, bead-bound DNA was also prepared and recovered as described above (Chapter 2). MMS damaged plasmid was used at 10 ng/μL in this chapter.

### 3.2.1 Additional *Xenopus* egg extract methods

To make LSS (modified from Blow, 1993), eggs were harvested and dejellied in 2.2% cysteine, pH 7.8 containing 1 mM ethylene glycol tetraacetic acid (EGTA; Sigma). Eggs were washed three times in Barth solution (88 mM NaCl, 2 mM KCl, 1 mM $MgCl_2$, 0.5 mM $CaCl_2$, 15 mM Tris pH 7.4) and activated with the addition of the calcium ionophore A23187 (2 μg/mL; Sigma). After activation, A23187 was washed away with additional washes in Barth solution. Eggs were then washed thoroughly in chilled extraction buffer (50 mM KCl, 5 mM $MgCl_2$, 2 mM DTT, 50 mM sucrose, 50 mM HEPES-KOH pH 7.6) and packed via centrifugation at 1100 rpm in an IEC CL2 tabletop centrifuge. Crude extract was then prepared by centrifugation in a Sorvall SB6 rotor as for crude S phase extract and supplemented with cycloheximide (50 μg/mL), DTT (1 mM), aprotinin (10 μg/mL), leupeptin (10 μg/mL), and cytochalasin B (2.5 μg/mL). This extract was diluted 10-15% with extract dilution buffer (EDBS; 50 mM KCl, 2 mM DTT, 0.4 mM $MgCl_2$, 0.4 mM EGTA, 10% sucrose, 50 mM HEPES-KOH, pH 7.6, 1 μg/mL aprotinin, 1 μg/mL leupeptin).

Extract was cleared by centrifugation at 30,000 rpm in an SW41 rotor (Beckman) for 40 minutes. Lipids were removed and LSS was harvested, leaving the mitochondrial layer behind. LSS was filtered through a nitex membrane and frozen in liquid nitrogen before storing at -80°C in single use aliquots.

### 3.2.2 Plasmid construction, mutagenesis, and protein purification

Human PCNA$^{SC}$ was generated in a multistep process. PCNA is highly conserved between *Xenopus* and humans, and human PCNA was previously used to study PCNA function in *Xenopus* egg extract (Arias and Walter, 2006; Havens and Walter, 2009; Havens et al., 2012). First, the PCNA gene (Arias and Walter, 2006), along with DNA encoding a Prescission Protease cleavage site 5' to the PCNA gene and the desired linker sequences 3' to the PCNA gene, was inserted into a pET28 vector using the primers 5'-AAAACATATGCTCGAAGTGTTGTTCCAAGGGCCCGAATTCGAGGCGCG CCTGGTCCAGG-3' and 5'-TTTTCTCGAGTTTGCTAGCTTTGGATCCGTTGGACTGC GAATTAGATCCTTCTTCATCCTCGATCTTGGG-3' paired with restriction enzyme digest with NdeI (NEB) and XhoI (NEB). Restriction enzyme recognition sites are underlined. The two linker sequences were nearly identical on the amino acid level (peptide linker #1=NSQ*SNG*SGA, peptide linker #2=NSQ*ASN*SGA, differences are italicized). These linker sequences were chosen based on previous construction of yeast single chain PCNA (McNally et al., 2010). DNA encoding the first linker sequence contains a BamHI (NEB) restriction enzyme recognition site, and that of the second linker site contains a NheI (NEB) recognition site (underlined in primer sequences). Second, individual wildtype PCNA or *pcna-79* genes were amplified from constructs previously published (Arias and Walter, 2006) with primers containing the corresponding restriction enzyme sites for insertion between the two linkers or between the linker and the stop codon: 5'-AAAAGGATCCGGAGCATTCGAGGCGCGCCTGGTC CAGGGC-3'

and 5'-TTTTGCTAGCCTGGCTGTTAGATCCTTCTTCATCCTCGAT CTTGGG-3' or 5'-AAAAGCTAGCAACAGTGGCGCATTCGAGGCGCGCCTGGTCC AGGGC-3' and 5'-TTTTCTCGAGGTATACCTAAGATCCTTCTTCATCC TCGATCTTGGG-3'. Again, restriction enzyme recognition sites are underlined. The second and third PCNA subunits were inserted through restriction enzyme digest and ligation. Importantly, any mutagenesis of an individual subunit must be performed on the monomeric PCNA construct, not in the context of $PCNA^{SC}$, since the PCNA-coding region of each subunit within the $PCNA^{SC}$ construct is identical.

$PCNA^{SC}$ expression was induced at 18°C overnight in BL21(DE3) cells with 0.5 mM IPTG. Cells were lysed via sonication in lysis buffer containing 300 mM NaCl, 10% sucrose, 5 mM imidazole, 5 mM β-mercaptoethanol, 50 mM Tris, pH 8, and protease inhibitors. After clearing the lysate at 30,000xg, the soluble fraction of $PCNA^{SC}$ was bound to Ni-NTA beads for 1 hour. Beads were washed with lysis buffer containing 20 mM imidazole and applied to an empty mini protein purification column (Biorad). Beads were additionally washed with 30 volumes of lysis buffer containing 20 mM imidazole and then eluted in buffer containing 500 mM imidazole and protease inhibitors 0.5 mM PMSF, 1 mM benzamidine, 10 µg/mL aprotinin, 10 µg/mL leupeptin. Elution fractions containing the most $PCNA^{SC}$ were combined. The His tag was cleaved with Prescission Protease during dialysis into buffer containing 50 mM NaCl, 10% glycerol, 1 mM DTT, 1 mM EDTA, and 50 mM Tris, pH 7.5. The presence of protease inhibitors did not prevent His tag cleavage by Prescission Protease. The resulting protein preparation was concentrated using concentrator columns (Millipore). In extract, $PCNA^{SC}$ was added to a final concentration of 150 ng/µL (5 µM each subunit) unless otherwise indicated.

Expression and purification of monomeric PCNA has been previously described (Arias and Walter, 2006): PCNA expression was induced at 37°C for 4 hours in BL21(DE3)

cells with 0.5 mM IPTG. Cells were lysed via sonication in the same lysis buffer used for $PCNA^{SC}$. After clearing the lysate at 30,000xg, the soluble fraction of PCNA was bound to Ni-NTA beads for 1 hour. Beads were washed with lysis buffer containing 50 mM imidazole and applied to an empty mini protein purification column (Biorad). Beads were additionally washed with 30 volumes of lysis buffer containing 100 mM imidazole and then eluted in lysis buffer containing 500 mM imidazole. The His tag was cleaved with Prescission Protease during dialysis into buffer containing 50 mM NaCl, 10% glycerol, 1 mM DTT, 1 mM EDTA, and 50 mM Tris, pH 7.5. The resulting protein preparation was concentrated using concentrator columns (Millipore).

$Cdt1^{1-243}$ was generated previously by Dr. Courtney Havens (Havens and Walter, 2009).

### 3.2.3 Immunological methods

Antibodies for immunoblotting (Human/*Xenopus* PCNA and *Xenopus* Cdt1, Cdt2, Orc2, and MCM7) are described above (Chapter 2). The GST antibody was purchased from Santa Cruz. A previously described polyclonal PCNA antibody (Kochaniak et al., 2009) was used to immunodeplete endogenous PCNA from HSS or LSS. Depletions were performed using a 3:1 ratio of PCNA antibody prebound to Protein A Sepharose Fastflow resin (Amersham Biosciences). Antibody-bound resin was used at a 1:5 ratio of resin to egg extract. HSS was depleted in three one-hour rounds at 4°C, and LSS was depleted in two one-hour rounds at 4°C.

### 3.2.4 Recovery of His-tagged proteins from extract

His-tagged proteins and their interacting partners were recovered with Ni-NTA beads in interaction buffer (5 µL beads per 25 µL sample; 100 mM NaCl, 5 mM $MgCl_2$, 10% glycerol, 0.1% NP-40, 200 mM Tris, pH 8). Proteins were first incubated in extract or ELB for 30 minutes. Beads were bound to protein for 1 hour at 4°C and then washed

three times with interaction buffer containing increasing concentrations of imidazole (0, 5, and 20 mM). Bound proteins were eluted in 2x SDS-PAGE sample buffer.

### 3.2.5 DNA replication assays in *Xenopus* egg extract (HSS and LSS)

Replication assays were carried out according to the methods outlined for NPE (Chapter 2) with the exception that there was no separate licensing step. The DNA (15 ng/μL M13 mp18 ssDNA (NEB) for primer extension in HSS and demembranated sperm chromatin for replication in LSS) was added directly to extract containing $\alpha^{32}$P-dATP. Reactions were stopped at the desired time points, and the reaction products were visualized as described above (Chapter 2).

## *3.3 Results*

### 3.3.1 The role of the Cdt2 C-terminal PIP box during CRL4$^{Cdt2}$ function

Recent data suggest that Cdt2 contacts more than one PCNA subunit (Kim et al., 2010). Indeed, Dr. Courtney Havens, a previous post-doctoral fellow in the Walter lab, identified a highly conserved PIP box in the C-terminus of Cdt2 (Figure 3.1). To determine whether Cdt2 must interact with an unoccupied subunit of PCNA in order to target its substrates for destruction, we took two different approaches: Havens mutated the PIP box of Cdt2 and determined its effect on Cdt1 destruction, and I generated a single chain PCNA trimer (PCNA$^{SC}$) that contains one wildtype subunit and two IDCL mutant subunits connected through a flexible linker.

Mutation of the putative Cdt2 PIP box had minimal effects on Cdt1 or Set8 ubiquitylation and destruction (Cdt1 destruction shown in Figure 3.2A, ubiquitylation shown in Figure 3.2B), but it had a substantial effect on stable Cdt2 recruitment to chromatin (Figure 3.2B, compare to C). These data suggest that CRL4$^{Cdt2}$ function does not require direct contact between the Cdt2 PIP box and the IDCL of a PCNA subunit despite improved

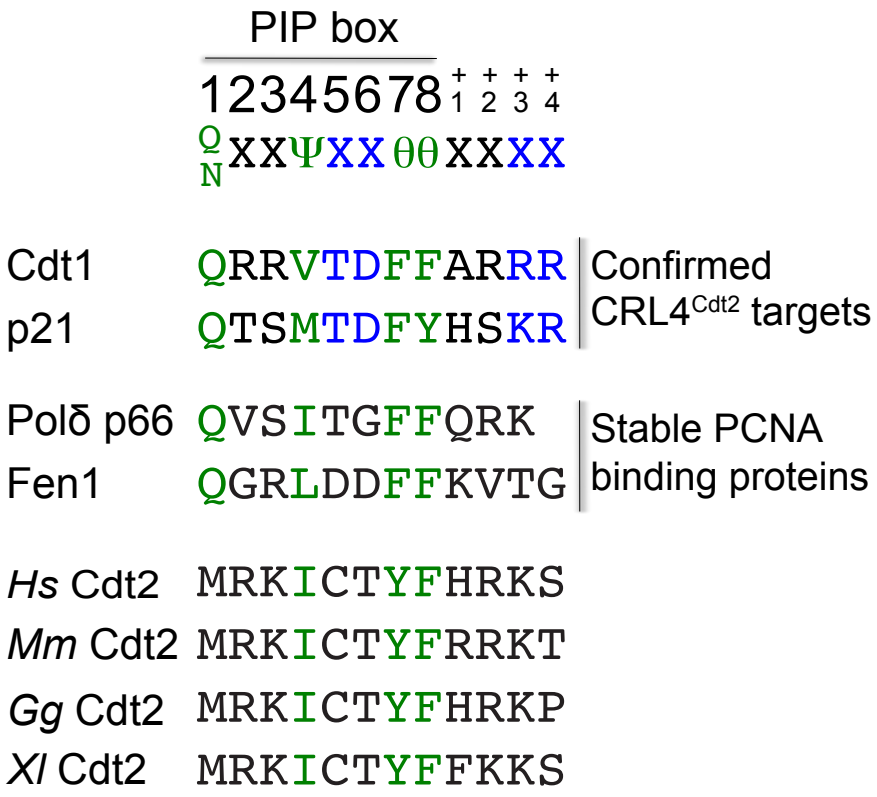



**Figure 3.1: Cdt2 contains a PIP box in its C-terminus.** The conserved PIP box and PIP degron residues are shown at the top. The putative C-terminal PIP box of Cdt2 is shown across several species and compared to the human PIP boxes of confirmed $CRL4^{Cdt2}$ targets and stable PCNA binding proteins. Ψ represents any hydrophobic residue (I/L/V/M), and θ represents any aromatic residue (Y/F/W). Residues shown in green are part of the conserved PIP box, and residues shown in blue additionally contribute to the "PIP degron."

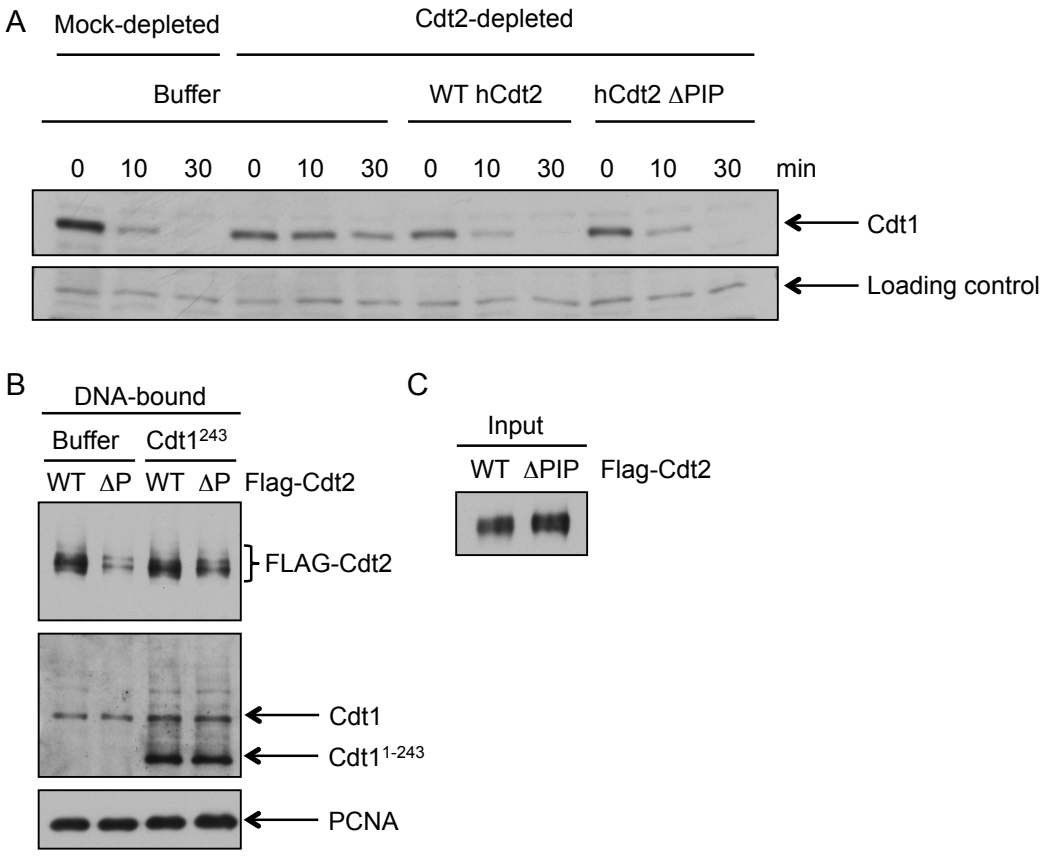



**Figure 3.2: Mutation of Cdt2's putative PIP box results in decreased Cdt2 association with chromatin but does not affect substrate destruction.** The data in this figure was generated by Dr. Courtney Havens, a previous post-doctoral fellow in the Walter lab. **A)** Cdt2 was immunodepleted and replaced with wildtype Flag-Cdt2 or Flag-Cdt2 ΔPIP. Bulk Cdt1 destruction was monitored. **B)** Chromatin-associated Cdt2 and Cdt1 are shown. $Cdt1^{1-243}$ was optionally added because it does not associate with chromatin through Orc. Its association with depends only on association with PCNA. **C)** Cdt2 in total extract for the experiment shown in (B).

PIP box-dependent Cdt2 association, at least for the well-characterized substrates Cdt1 and Set8. However, the secondary contact could be necessary for destruction of weaker substrates that have not been tested. Also, this mutation analysis alone may not reveal whether Cdt2 associates with a PCNA subunit that is distinct from the substrate-bound PCNA subunit, as mutation of the putative PIP box might not be sufficient to eliminate interaction with PCNA. Because mutation of the Cdt2 PIP box affected stable Cdt2 recruitment to chromatin but did not affect substrate ubiquitylation or destruction, this result is consistent with some direct Cdt2 contact with PCNA. However, the biological relevance of this contact is not clear and may not be related to $CRL4^{Cdt2}$ function.

### 3.3.2 Characterization of a single chain PCNA polypeptide

I generated a single chain PCNA trimer ($PCNA^{SC}$) in which the three subunits of PCNA are fused into a single polypeptide via flexible sequences of amino acids (McNally et al., 2010) (Figure 3.3). I then mutated the IDCL (depicted in detail in Figure 1.2, shown as a notch in Figure 3.3A) in one, two, or three subunits to establish how many PIP box interaction sites are required for PCNA function during $CRL4^{Cdt2}$-dependent ubiquitylation and other PCNA-dependent processes. Changing the number of PIP box binding sites on each PCNA trimer allows us to ask whether an IDCL other than the one contacted by the substrate is required for $CRL4^{Cdt2}$ activity. The IDCL mutation, designated *pcna-79*, eliminates interaction between PCNA and multiple conserved PIP box residues, and it inhibits DNA Polymerase δ binding in yeast (Eissenberg et al., 1997). Here, I will refer to each individual subunit within $PCNA^{SC}$ as either wildtype (WT) or *pcna-79* (79) (Figure 3.3A). If Cdt2 or another protein must bind to a separate IDCL on PCNA for $CRL4^{Cdt2}$ function, then neither $PCNA^{SC}$(79-79-79) nor $PCNA^{SC}$(wt-79-79) should support Cdt1 destruction because they do not allow multiple PIP box-containing proteins to bind at the same time.

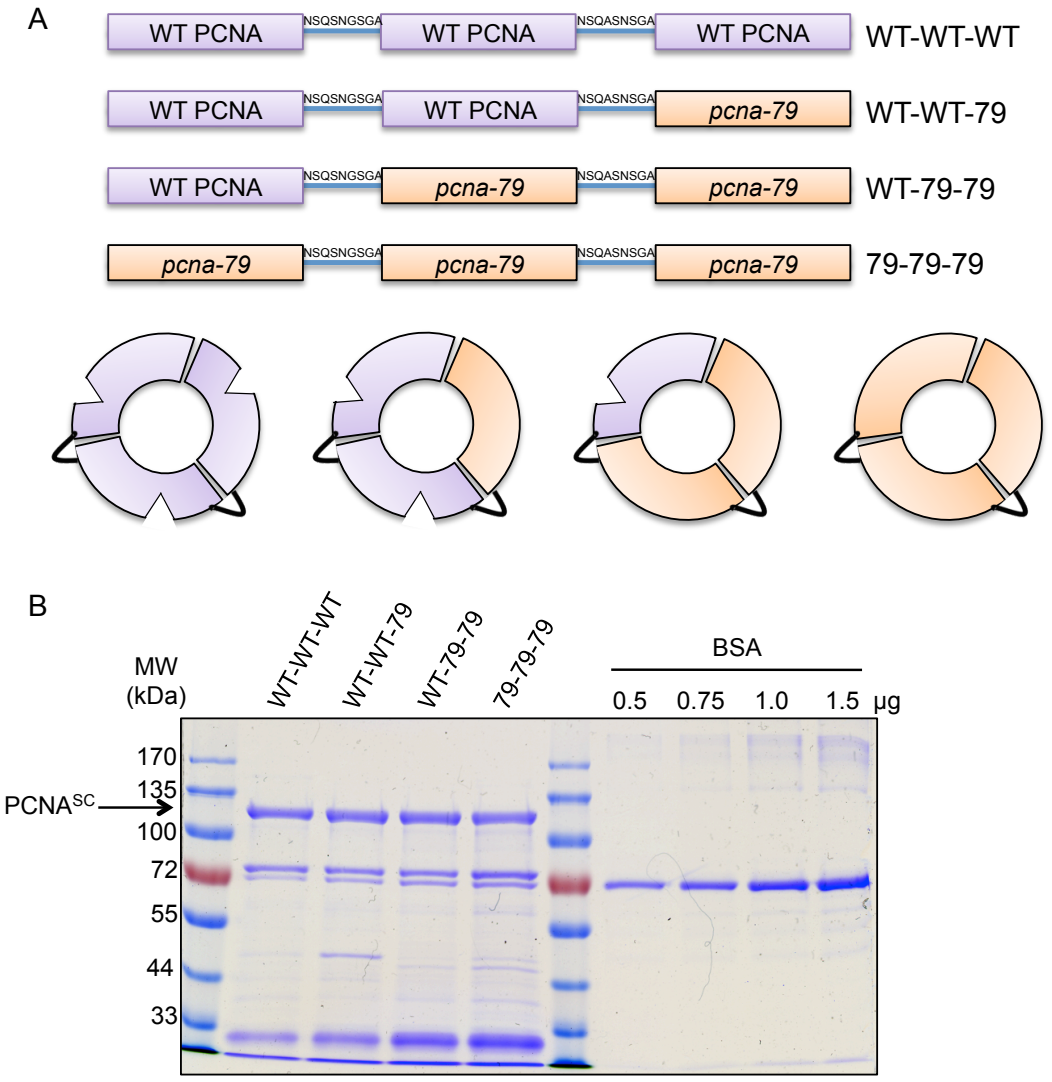


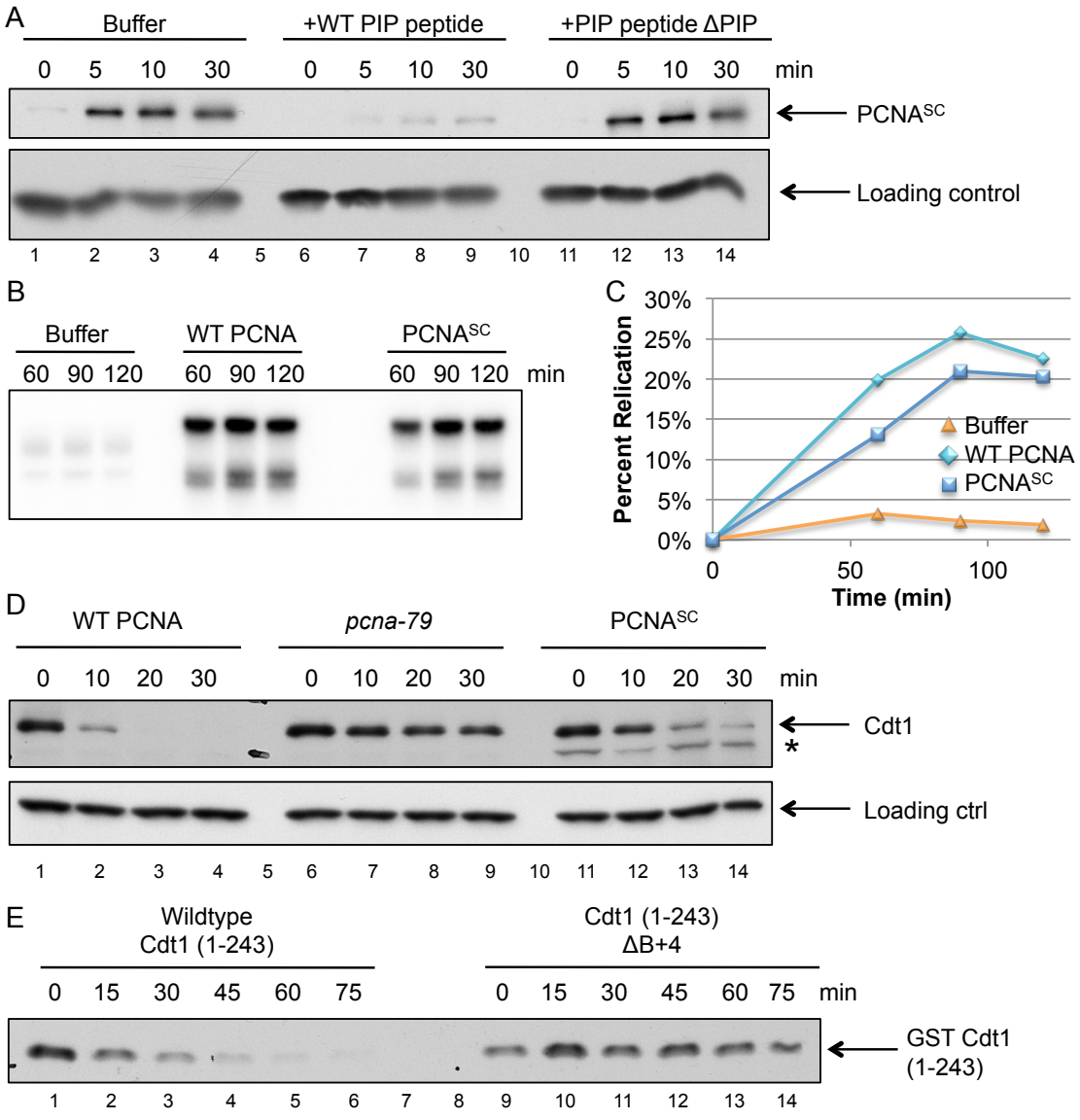

**Figure 3.3: PCNA$^{SC}$ constructs used in this work. A)** Three PCNA monomers were attached through flexible linker sequences described in Materials and Methods (3.2.2) and shown here. One, two or three PCNA subunits were mutated to eliminate their ability to interact with PIP box-containing proteins. The mutation was designated *pcna-79* based on the mutagenesis screen that identified it (Eissenberg et al., 1997). A functional IDCL is portrayed as a notch in a single PCNA subunit. **B)** Individual PCNA$^{SC}$ preparations. PCNA$^{SC}$ contains a number of non-specific contaminants that are not detected by the PCNA antibody.

I first characterized $PCNA^{SC}$ in our extract systems. To study $CRL4^{Cdt2}$-dependent destruction, we usually use HSS (see Chapter 1). When supplemented with methylated DNA damaged by guanine methylation, HSS supports nucleotide excision repair (NER), the last step of which is PCNA-dependent gap filling. Thus, addition of methylated (MMS-damaged) DNA leads to $CRL4^{Cdt2}$-dependent destruction of its substrates (Jin et al., 2006). To characterize $PCNA^{SC}$ in HSS, I immunodepleted endogenous PCNA and replaced it with $PCNA^{SC}$. I observed that $PCNA^{SC}$ binds chromatin (Figure 3.4A, lanes 1-4), and an excess of the PIP box peptide of p21 (described in Chapter 2) inhibited loading (Figure 3.4, lanes 5-8). This outcome was expected because the PCNA clamp loader, replication factor C (RFC), must bind the IDCL to load PCNA onto DNA through its PIP box (Oku et al., 1998). A PIP box peptide in which the conserved PIP box residues have been mutated to alanine did not affect $PCNA^{SC}$ association with chromatin (Figure 3.4, lanes 9-12).

To study PCNA-dependent DNA replication in HSS, we used a primer extension assay. In this assay, Polα/primase primes a single-stranded plasmid template, and the replicative polymerases Polδ and Polε extend the primer in a PCNA-dependent manner, much like leading strand synthesis during chromosomal replication. Both $PCNA^{SC}$ and wildtype PCNA supported this type of replication in HSS (Figure 3.4B, C). Most importantly for studying $CRL4^{Cdt2}$-dependent destruction, $PCNA^{SC}$ supported destruction of Cdt1 in the presence of damaged plasmid or damaged linear DNA, though it was somewhat slower than destruction triggered by wildtype PCNA (Figure 3.4D, compare lanes 11-14 to lanes 1-4). Finally, to confirm that in the presence of $PCNA^{SC}$ the destruction of Cdt1 followed the established rules for $CRL4^{Cdt2}$-dependent proteolysis, I compared destruction of wildtype Cdt1 to Cdt1 lacking the critical B+4 residue and saw stabilization of the mutant, as expected (Figure 3.4E). Together, these data show that

**Figure 3.4: PCNA$^{SC}$ is functional in HSS. A)** PCNA$^{SC}$ associates with DNA in an IDCL-dependent manner. PIP box peptide was added to a final concentration of 100 μM. Chromatin spin downs were performed with 250 mM KCl. **B)** PCNA$^{SC}$ rescues PCNA-dependent primer extension on a single-stranded plasmid template in HSS. Replication of a single-stranded M13 plasmid was monitored. **C)** Quantification of (B). **D)** PCNA$^{SC}$ promotes CRL4$^{Cdt2}$ function as measured by Cdt1 destruction, albeit slower than wildtype monomeric PCNA. The asterisk (*) indicates a non-specific band due to contaminants in the PCNA$^{SC}$ prep. An extract-specific, non-Cdt1 band detected by the Cdt1 antibody is shown as a loading control. **E)** Cdt1 destruction triggered by PCNA$^{SC}$ depends on a PIP degron. Cdt1$^{1-243}$ cannot associate with ORC and is therefore completely dependent on PCNA for its association with DNA.

**Figure 3.4** (continued)

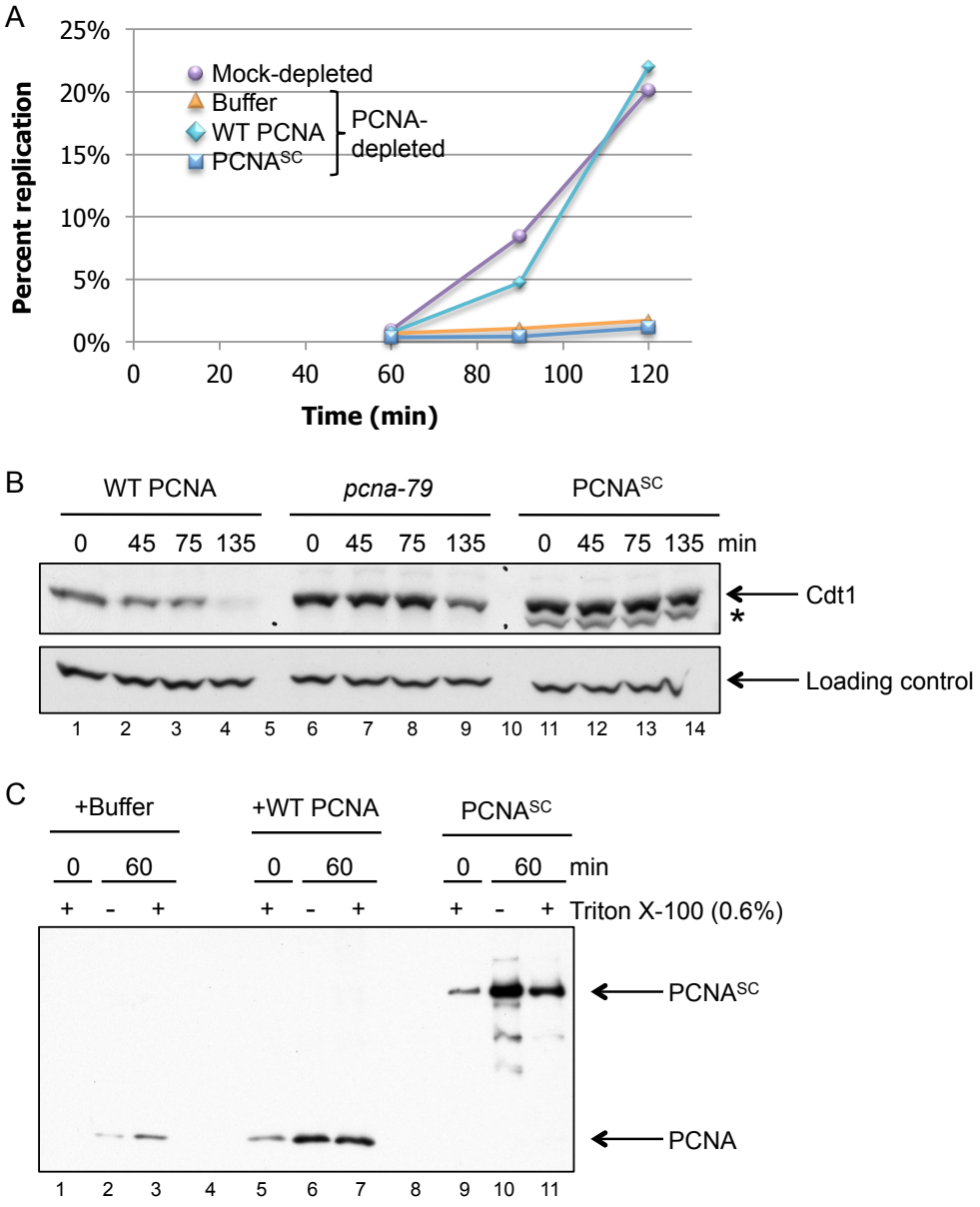


PCNA$^{SC}$ can replace endogenous PCNA in HSS, and that we can use PCNA$^{SC}$ to study a number of PCNA-related events in HSS, including CRL4$^{Cdt2}$-dependent protein destruction.

I next tested whether PCNA$^{SC}$ can substitute for wildtype PCNA, assembled of three individual PCNA monomers, during chromosomal replication and Cdt1 destruction in LSS. However, after depleting endogenous PCNA from LSS, I could not rescue chromosomal DNA replication or replication-dependent Cdt1 destruction under conditions where wildtype monomeric PCNA rescued both (Figure 3.5A, B, compare lanes 12-15 to lanes 1-4). One possible explanation for the failure of PCNA$^{SC}$ to support chromosomal replication and Cdt1 destruction in LSS is that unlike wildtype PCNA, PCNA$^{SC}$ fails to enter the nucleus. Though the linker sequences in PCNA$^{SC}$ did not alter the PCNA nuclear localization sequence (NLS) (Kim and Lee, 2012), PCNA$^{SC}$ is much larger than each PCNA monomer and might therefore require attachment of a different NLS to function in LSS. First, I looked at whether PCNA$^{SC}$ could enter the nucleus and interact with sperm chromatin. I was able to recover PCNA$^{SC}$ from the chromatin-associated fraction in a traditional chromatin spin down experiment (Figure 3.5C, lane 11). To ensure this association did not take place during the spin down itself, I performed two important controls: 1) I performed a chromatin spin down at the start of the reaction, before nuclei are expected to form, and found that PCNA$^{SC}$ did not associate noticeably with sperm chromatin (Figure 3.5C, lane 9); and 2) I omitted Triton from the spin down to prevent permeabilization of the nucleus, and I still recovered PCNA$^{SC}$ (Figure 3.5C, lane 10). These results suggest that PCNA$^{SC}$ can enter the nucleus and associate with sperm chromatin. In the presence of PCNA$^{SC}$, the association between endogenous PCNA and chromatin also became undetectable, suggesting that PCNA$^{SC}$ competes with endogenous PCNA for chromatin loading (Figure 3.5C, compare lanes 9-11 to lanes 1-3).

**Figure 3.5: PCNA$^{SC}$ does not rescue PCNA depletion in LSS. A)** Wildtype monomeric PCNA but not PCNA$^{SC}$ supports chromosomal DNA replication in LSS. **B)** Wildtype monomeric PCNA, but not PCNA$^{SC}$ supports Cdt1 destruction during replication in LSS. An extract-specific, non-Cdt1 band detected by the Cdt1 antibody is shown as a loading control. **C)** PCNA$^{SC}$ associates with sperm chromatin in LSS. Triton X-100 was omitted in some samples to compare the total nuclear PCNA$^{SC}$ (no Triton) with chromatin-bound PCNA (+0.6% Triton). 250 mM KCl was used to recover sperm chromatin. Detection of PCNA$^{SC}$ in the recovered fraction did not depend on membrane permeabilization by Triton.

**Figure 3.5** (continued)



Several other control experiments would be necessary to confirm this result, but based on this single experiment, it appears that there are no defects in the ability of $PCNA^{SC}$ to enter the nucleus or to load onto sperm chromatin. Therefore, $PCNA^{SC}$ cannot function in LSS, but the reason for this failure is yet unknown. To test whether $PCNA^{SC}$ can support chromosomal replication and Cdt1 destruction in the absence of a nucleus, we would ideally test $PCNA^{SC}$ function in NPE. However, this is not yet possible because the levels of endogenous PCNA are so high in NPE that immunodepletion is not possible.

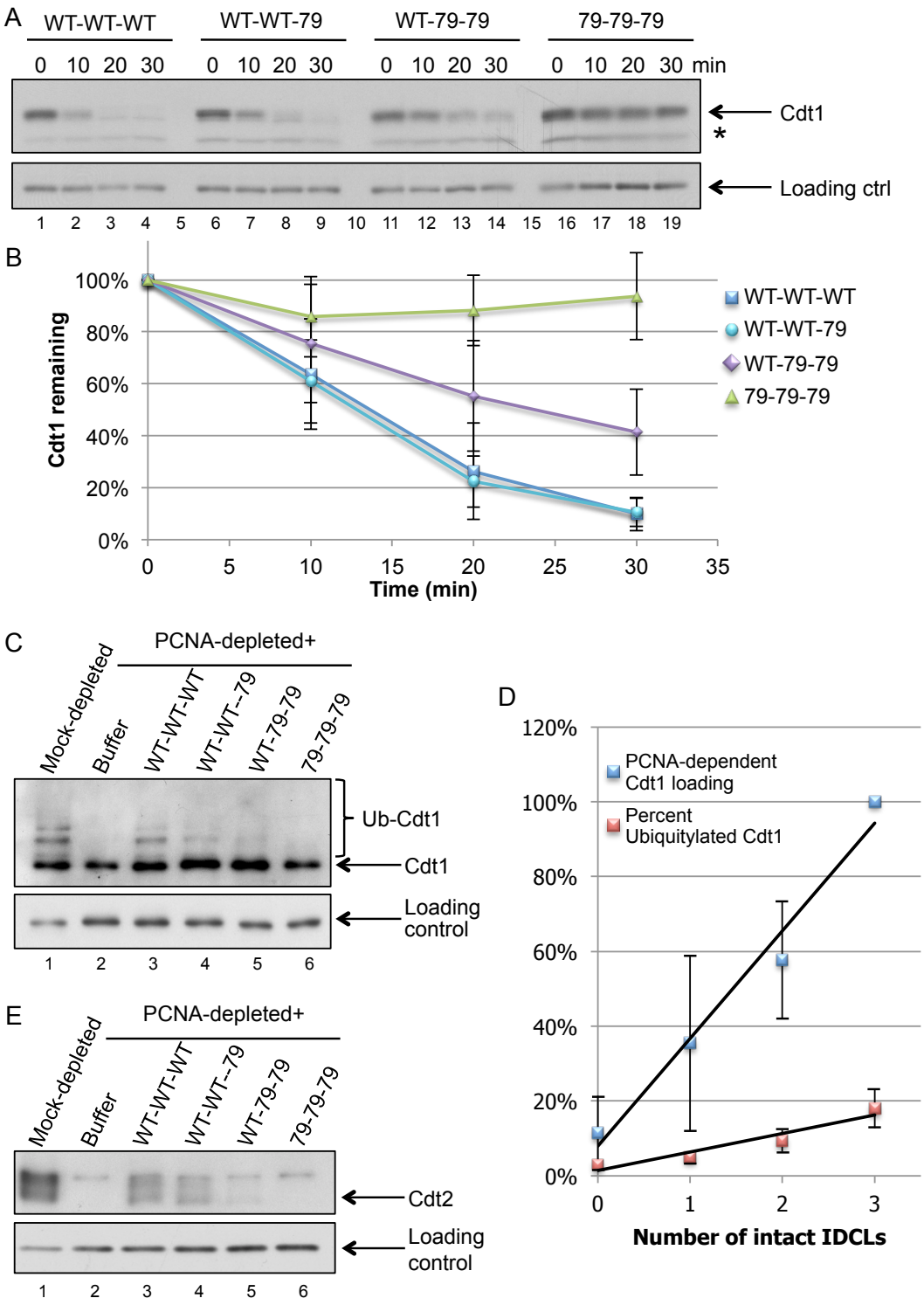
### 3.3.3 The role of multiple interdomain connector loops in $CRL4^{Cdt2}$ function

I next generated $PCNA^{SC}$ constructs that contain two, one, or no wildtype subunits to determine how many intact IDCLs within an individual PCNA trimer are required in HSS for $CRL4^{Cdt2}$-dependent destruction, Cdt2 recruitment to chromatin, substrate ubiquitylation, and substrate loading to chromatin. As expected, $PCNA^{SC}$(79-79-79), which consists of only mutant IDCL subunits, did not support Cdt2 recruitment or Cdt1 destruction (Figure 3.6A-D; lanes 16-18 in A, lane 6 in D) because Cdt1 must bind PCNA through its own PIP box. Also as expected, $PCNA^{SC}$(WT-WT-79) supported complete Cdt1 destruction with similar kinetics to destruction triggered by $PCNA^{SC}$(WT-WT-WT), indicating that $CRL4^{Cdt2}$ function does not require all three intact PIP box binding sites (Figure 3.6A, B; compare lanes 6-9 to lanes 1-4 in A). $PCNA^{SC}$(WT-79-79) also supported Cdt1 destruction, but at a reduced rate compared to $PCNA^{SC}$(WT-WT-WT) and $PCNA^{SC}$(WT-WT-79) (Figure 3.6A, B; compare lanes 11-14 to lanes 1-4 and 6-9 in A), indicating that binding of multiple PIP box proteins to a single PCNA trimer is not absolutely required for all $CRL4^{Cdt2}$ function, but it may enhance destruction.

I wanted to address why PCNA containing only one intact IDCL supported only partial Cdt1 destruction. Therefore, I considered each step involved in $CRL4^{Cdt2}$ function,

**Figure 3.6: A single IDCL on PCNA is sufficient for partial CRL4$^{Cdt2}$ function. A)** PCNA-depleted HSS was rescued with PCNA$^{SC}$ containing mutated IDCLs on varying numbers of subunits. PCNA$^{SC}$ containing only a single wildtype subunit supported partial Cdt1 destruction. An extract-specific, non-Cdt1 band detected by the Cdt1 antibody is shown as a loading control. **B)** Quantification of (A). Cdt1 bands were quantified using Image J software. Each Cdt1-specific band was normalized to a loading control, and the starting levels of Cdt1 were set to 100%. N≥5 for all samples (N=10 for WT-WT-WT, N=7 for WT-WT-79, N=6 for WT-79-79, N=5 for 79-79-79). **C)** Decreasing the number of IDCLs on PCNA leads to decreased Cdt1 binding and ubiquitylation on chromatin. Streptavidin is shown as a loading control. **D)** Quantification of (C). PCNA-dependent Cdt1 binding is reported as the difference between total Cdt1 signal and Orc-dependent Cdt1 signal (buffer control); Cdt1 binding stimulated by PCNA$^{SC}$(WT-WT-WT) is set to 100%. The percent of ubiquitylated Cdt1 is reported as the ratio of ubiquitylated to total Cdt1. **E)** Decreasing the number of IDCLs on PCNA leads to decreased Cdt2 recruitment. Chromatin spin downs were performed with 100 mM KCl. Orc2 was used as a loading control.

**Figure 3.6** (continued)

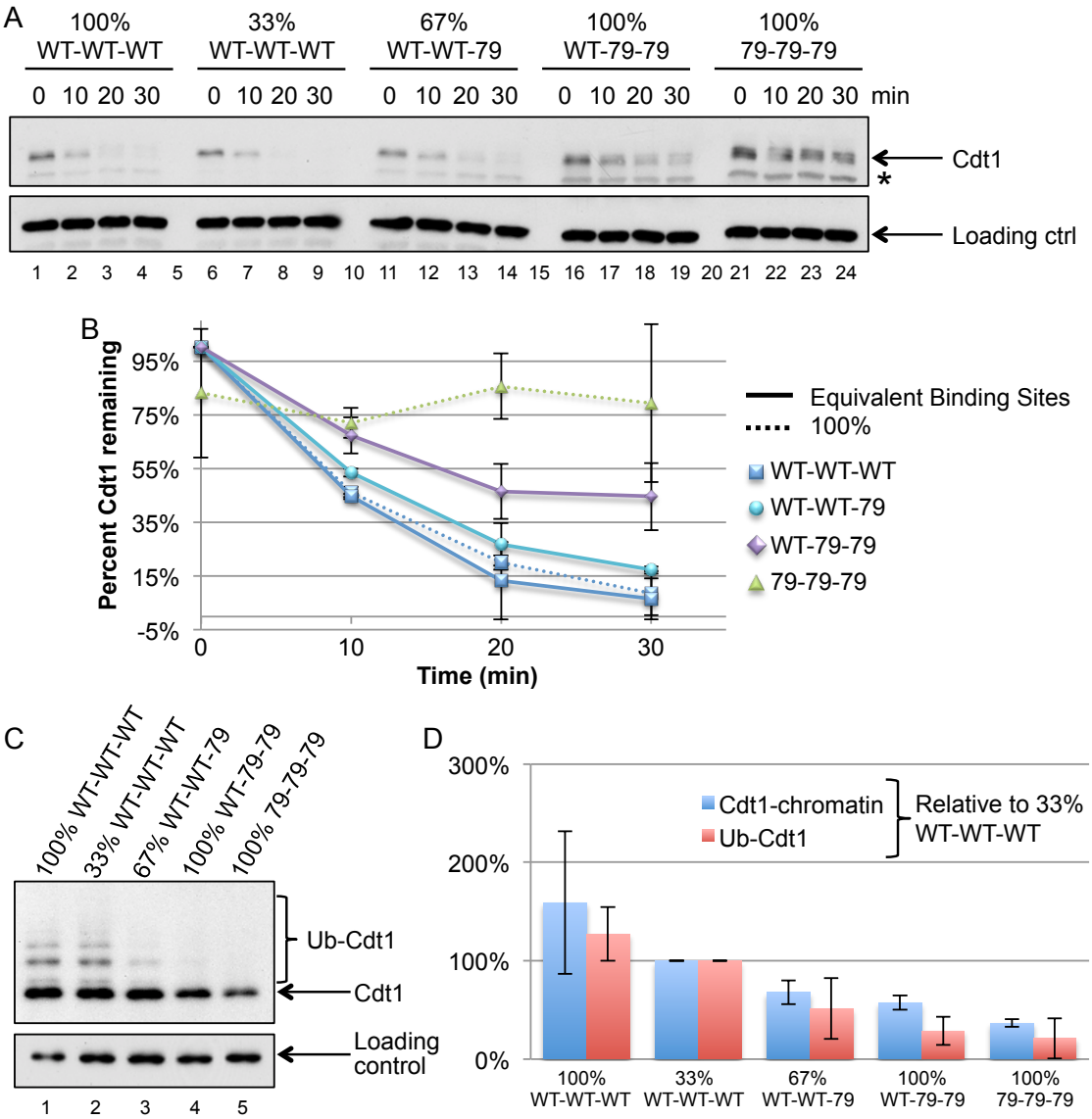


starting with substrate binding. I found that decreasing the number of wildtype IDCLs by mutation of only one subunit in PCNA$^{SC}$(WT-WT-79) led to decreased Cdt1 recruitment and ubiquitylation (Figure 3.6C, compare lanes 3 and 4, see D for quantification); recruitment and ubiquitylation in extract containing PCNA$^{SC}$(WT-79-79) was even lower (Figure 3.6C (lane 5), D). Cdt2 recruitment to chromatin followed the same trend (Figure 3.6E). It is therefore plausible that defects in Cdt1 destruction in the presence of PCNA$^{SC}$(WT-79-79) might simply be due to a decreased number of binding sites for CRL4$^{Cdt2}$ substrates rather than an intrinsic requirement for at least two binding sites. Under these conditions, defects in bulk destruction might only become evident when ligase recruitment and substrate ubiquitylation are below a certain threshold.

To distinguish the possibility that decreased Cdt1 proteolysis in the presence of PCNA$^{SC}$(WT-79-79) is due to decreased Cdt1 binding sites from the possibility that a second available PIP box binding site contributes to CRL4$^{Cdt2}$ function, I repeated the destruction assays and monitored chromatin-bound Cdt1 and Cdt1 ubiquitylation in extract containing 1/3 the amount of PCNA$^{SC}$(WT-WT-WT) relative to PCNA$^{SC}$(WT-79-79). Under these conditions, the molarity of intact IDCLs was the same in both reactions. Even after equalizing the molarity of intact IDCLs, PCNA$^{SC}$(WT-WT-WT) still supported more efficient Cdt1 destruction than PCNA$^{SC}$(WT-79-79) (Figure 3.7A, B; compare lanes 16-19 to lanes 6-9 in A). Decreasing the total amount of PCNA$^{SC}$(WT-WT-WT) in extract did have a modest effect on Cdt1 binding and ubiquitylation, but equalizing the number of wildtype subunits in extract also did not restore decreased Cdt1 binding and ubiquitylation caused by mutation of additional IDCLs (Figure 3.7C, D; compare lanes 2-4 in C). However, the amount of PCNA that *loads* onto DNA relative to total PCNA in solution is very low. Though a *pcna-79* mutation does not substantially affect RFC-dependent PCNA loading *in vitro* (Eissenberg et al., 1997), RFC makes several contacts

**Figure 3.7: Normalizing the number of available IDCLs in total extract does not resolve the partial rescue problem.** A) Cdt1 destruction supported by indicated amounts of $PCNA^{SC}$ relative to the amount used in Figure 3.6. An extract-specific, non-Cdt1 band detected by the Cdt1 antibody is shown as a loading control. **B)** Quantification of (A). The average of two experiments is shown; error bars represent the standard deviation. **C)** Normalizing the number of available IDCLs does not restore equal Cdt1 loading or ubiquitylation. Orc2 is shown a loading control. **D)** Quantification of (C). Three separate experiments were averaged, and error bars represent the standard deviation.

**Figure 3.7** (continued)



with the IDCL (Oku et al., 1998), so it is possible that the number of wildtype IDCLs on DNA was still dissimilar in the two samples because of differences in $PCNA^{SC}$ loading by RFC (Oku et al., 1998). Thus, I calculated how much of each $PCNA^{SC}$ construct is necessary in total extract to generate a three-fold difference in the DNA-bound concentration of $PCNA^{SC}$(WT-79-79) relative to $PCNA^{SC}$(WT-WT-WT) (Figure 3.8A). When I added 2.5% of $PCNA^{SC}$(WT-WT-WT) and 33% of $PCNA^{SC}$(WT-79-79) relative to the amount used in the original destruction assays (150 ng/μL, Figure 3.6), the levels of wildtype IDCL subunits of $PCNA^{SC}$ on chromatin were similar (Figure 3.8B, lanes 1-3 contain the same sample at different dilutions, as do lanes 5-7). This low concentration of $PCNA^{SC}$(WT-WT-WT) did not cause defects in Cdt1 destruction like $PCNA^{SC}$(WT-79-79), despite a level of chromatin-bound $PCNA^{SC}$ that should correspond to an equivalent number of binding sites for PIP box containing proteins (Figure 3.8C, compare lanes 11-14 to lanes 7-10). Hence, having a second IDCL present on the same PCNA trimer to which the substrate can bind seems to somehow facilitate $CRL4^{Cdt2}$ function.

Nonetheless, the observation that normalizing the amount of chromatin-bound wildtype PCNA subunits did not resolve differences in destruction could still be subject to multiple interpretations. If, for example, the $PCNA^{SC}$(WT-WT-WT) protein preparation contains contaminating monomeric PCNA fragments not present in the $PCNA^{SC}$(WT-WT-79) preparation, it could allow for a higher concentration of wildtype PCNA subunits on DNA than those contributed by $PCNA^{SC}$. Upon long exposure of the DNA-bound PCNA blot, I detected a faint band migrating at the expected size of monomeric PCNA (~35 kDa) in samples containing $PCNA^{SC}$(WT-WT-WT) that was not visible in samples containing $PCNA^{SC}$(WT-79-79) (Figure 3.8B, compare lanes 1-3 to lanes 5-7 in the middle panel), suggestive of contamination by monomeric PCNA molecules due to proteolysis of the linker regions or breakdown of $PCNA^{SC}$ in the extract. This band was not visible in earlier

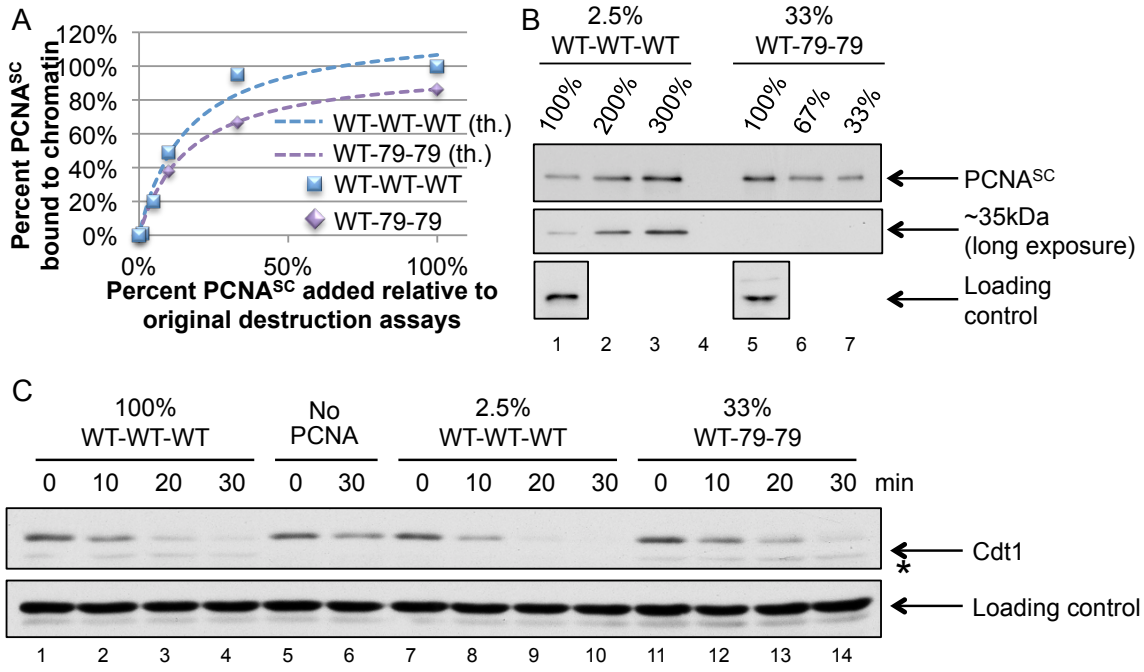



**Figure 3.8: Normalizing the number of available IDCLs on chromatin does not resolve the partial rescue problem. A)** PCNA$^{SC}$ was titrated into extract to generate a binding curve for PCNA$^{SC}$ loading to chromatin. Raw data points are shown along with the theoretical binding curve based on these data. **B)** Wildtype IDCLs on PCNA$^{SC}$ were normalized on chromatin. Chromatin-bound samples were titrated to allow careful quantification. 100% of the PCNA$^{SC}$(WT-WT-WT) sample should contain 1/3 the amount of bound PCNA$^{SC}$ as the 100% PCNA$^{SC}$(WT-79-79). MCM7 is shown as a loading control. **C)** Normalizing bound wildtype IDCLs does not restore equal Cdt1 destruction supported by PCNA$^{SC}$(WT-WT-WT) and PCNA$^{SC}$(WT-79-79). No quantification is shown because this is a single experiment.

experiments using $PCNA^{SC}$ prepared at a different time from the $PCNA^{SC}$ used in Figures 3.4-3.7 (data not shown).

Alternatively, $CRL4^{Cdt2}$ function could absolutely require multiple wildtype IDCLs on PCNA, but the single wildtype subunit of multiple $PCNA^{SC}$(WT-79-79) trimers could cooperate with the wildtype subunit on other trimers, stimulating the modest Cdt1 destruction I observe. Two trimers could interact in *trans*, with Cdt2 binding to the single wildtype IDCL on one trimer and the substrate binding to the other trimer (Figure 3.9A), or two trimers could interact in *cis*, with multiple $PCNA^{SC}$ polypeptides interacting to form a single trimer with additional attached subunits (Figure 3.9B). This second model seems unlikely for steric reasons but is still formally possible. In support of a potential interaction between $PCNA^{SC}$ and other PCNA molecules, I found that His-tagged PCNA trimers pulled down $PCNA^{SC}$ (Figure 3.9C, lane 9) even after $PCNA^{SC}$ has had the His tag removed through Prescission Protease. This experiment needs to be repeated to be conclusive and to confirm that Ni-NTA beads do not recover $PCNA^{SC}$ in the absence of His-tagged PCNA.

Wildtype PCNA homotrimers also reassembled in extract: when His-tagged homotrimers were mixed with extract, the His-tagged PCNA pulled out endogenous PCNA, suggesting that individual timers are not stable and that subunits can exchange or interact (Figure 3.9B, lane 5). The same was true for a mixture of His-tagged PCNA and untagged PCNA in ELB (Figure 3.9B, lane 7).

Collectively, these results suggest that multiple subunits on a single PCNA trimer may contribute to $CRL4^{Cdt2}$ function but are not absolutely required; however, I was not able to exclude a number of alternative explanations for our data, such as contaminating monomeric PCNA in $PCNA^{SC}$ protein preparations or the possibility that wildtype

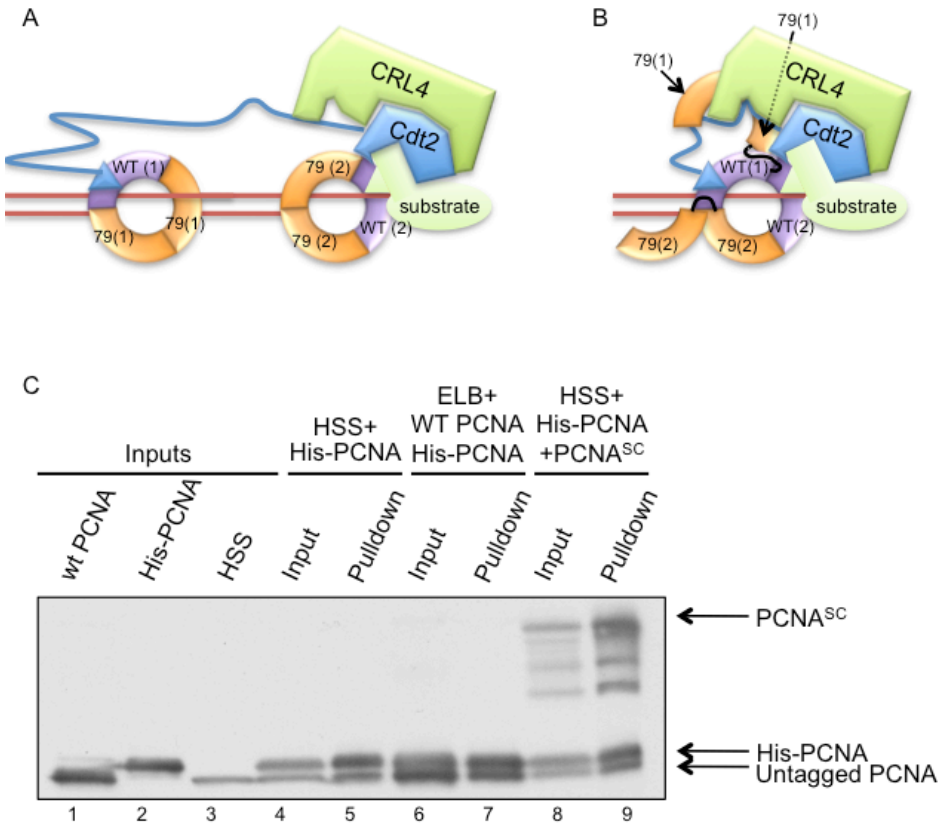



**Figure 3.9: Alternate explanations for the differences between PCNA$^{SC}$(WT-WT-WT) and PCNA$^{SC}$(WT-79-79). A)** Multiple PCNA heterotrimers could interact in *trans* to support Cdt2 binding to a separate IDCL from the CRL4$^{Cdt2}$ substrate. **B)** Multiple PCNA$^{SC}$ polypeptides could interact in *cis* to generate a single PCNA ring that contains more than one wildtype IDCL to support Cdt2 binding to a separate IDCL from the CRL4$^{Cdt2}$ substrate. **C)** PCNA$^{SC}$ interacts with monomeric PCNA in HSS and PCNA reassembles in extract. PCNA$^{SC}$ and His-PCNA were added to HSS for 30 minutes at room temperature before recovering His-PCNA with Ni-NTA beads.

subunits on different PCNA$^{SC}$(WT-79-79) heterotrimers cooperate to allow CRL4$^{Cdt2}$-dependent proteolysis.

### 3.3.4 A single IDCL on PCNA supports complete replication in HSS

To study the role of multiple IDCLs on PCNA during DNA replication, I used a primer extension assay on a single-stranded plasmid template. Though this assay does not address a role for multiple IDCLs in licensing, chromatin unwinding, or lagging strand synthesis, it does explore whether multiple PIP box containing proteins must bind a single PCNA trimer during leading strand elongation. Since PCNA$^{SC}$ can replace wildtype monomeric PCNA in such a replication assay (Figure 3.4B, C), I tested the effect of individual IDCL mutations in this assay. I immunodepleted PCNA and replaced it with each of the PCNA$^{SC}$ mutants. Even PCNA$^{SC}$(WT-79-79) stimulated full replication (Figure 3.10). Notably, PCNA$^{SC}$(79-79-79) supported a partial rescue of primer extension in this assay (Figure 3.10). In yeast and in *Xenopus* egg extract, *pcna-79* also supports a partial rescue of DNA replication on an M13 plasmid (Eissenberg et al., 1997; Havens, unpublished data), likely because this mutation does not completely abolish RFC-dependent PCNA loading and because primer extension depends primarily on the activity of Polε, which can interact with *pcna-79* almost as well as with wildtype PCNA (Eissenberg et al., 1997; Gomes and Burgers, 2000; Zhang et al., 1998).

In summary, my preliminary results suggest that neither CRL4$^{Cdt2}$-dependent proteolysis during DNA repair nor DNA replication in HSS strictly require multiple IDCLs on a single PCNA trimer, suggesting that at least for these two events, the “tool belt” model does not apply. However, further experimentation is necessary to resolve the described caveats. This does not mean these processes do not benefit from having adjacent IDCLs on PCNA. Indeed, a second IDCL seems to contribute to CRL4$^{Cdt2}$-dependent proteolysis, though it is unclear whether this contribution is simply due to the presence of additional

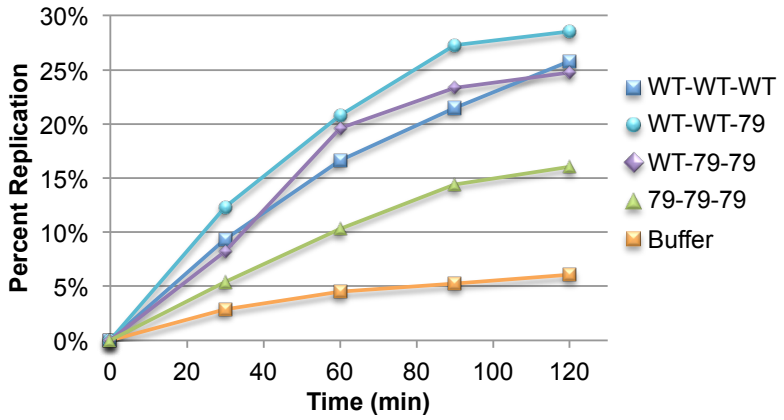



**Figure 3.10: PCNA$^{SC}$(WT-79-79) supports PCNA-dependent primer extension as well as PCNA$^{SC}$(WT-WT-WT).** PCNA$^{SC}$ was added to HSS containing single-stranded M13 plasmid DNA and radiolabeled dATP. Replication was monitored over time. Shown here is a quantification of the incorporation of radiolabeled dATP.

substrate binding sites because of the described stipulations. In conclusion, these data challenge a generally accepted model for PCNA function but require confirmation.

### *3.4 Discussion and Future Directions*

Until now, there has been no way to study the roles of individual subunits within the PCNA trimer. Though we can imagine PCNA as a "tool belt" on DNA, holding proteins poised for action should problems in replication arise (Freudenthal et al., 2011; Indiani et al., 2005; Sutton, 2010; Zhuang and Ai, 2010), heterotrimeric PCNA has not been available, so testing this model in eukaryotic systems has been difficult. I have generated a useful tool to study the contributions of distinct PCNA subunits to any biochemical process during S phase in *Xenopus* egg extract. I used $PCNA^{SC}$ to test the role of individual IDCLs in a single PCNA trimer during $CRL4^{Cdt2}$-dependent destruction and during primer extension in HSS. I found that only a single IDCL is necessary for primer extension on a single-stranded plasmid and for modest $CRL4^{Cdt2}$-dependent destruction of the well-characterized substrate Cdt1. However, $PCNA^{SC}$ containing only a single wildtype IDCL supported only partial Cdt1 destruction, and I was not able to conclusively determine whether this was simply due to decreased substrate binding or due to differences in ubiquitin ligase function.

In theory, $PCNA^{SC}$ should allow testing of a number of unanswered questions in the field of DNA replication: How many functional IDCLs does PCNA require to support full DNA replication? Does switching from a replicative polymerase to a translesion synthesis polymerase require multiple binding sites on PCNA? How do individual PCNA subunits contribute to DNA repair? How do PCNA modifications on individual subunits affect replication fork dynamics? Does PCNA unloading from DNA require protein interactions with multiple PCNA subunits? Does chromatin assembly require coordinated recruitment

of multiple factors to a single PCNA molecule? All of these processes affect the overall progression through S phase, and we have yet to understand the role of individual PCNA subunits. Unfortunately, addressing most of these questions requires either LSS or NPE: $PCNA^{SC}$ does not function in LSS (Figure 3.5), and we cannot immunodeplete endogenous PCNA from NPE since the levels are so high. Thus, so far it is impossible to replace endogenous PCNA with $PCNA^{SC}$ in NPE. Further studies utilizing $PCNA^{SC}$ will require fine-tuning to allow function in LSS, potentially by modifying the protein preparation to eliminate contaminants. Additionally, generating a protocol for successful depletion of PCNA from NPE would allow us to address almost any of these interesting questions. With additional effort, $PCNA^{SC}$ will be a useful tool for the future study of other PCNA-dependent events.

Finally, I was unable to resolve the difference in $CRL4^{Cdt2}$ function supported by $PCNA^{SC}$(WT-WT-WT) compared to $PCNA^{SC}$(WT-79-79). It remains possible that $CRL4^{Cdt2}$ requires multiple subunits of PCNA to function but that multiple $PCNA^{SC}$ molecules cooperate to facilitate substrate proteolysis. Confirmation of this phenomenon would introduce a new perspective on PCNA function that has not been seriously considered. Two reports from the Lee lab in Ontario have shown that two PCNA homotrimers can dimerize into a back-to-back doublet and that such dimerization is essential for concurrent binding of Polδ and chromatin assembly factor-1 (CAF1) (Naryzhny et al., 2006; 2005). The presence of such trimer dimers was independently confirmed using another method, but this study found that only a small fraction of PCNA molecules existed in such a configuration (De Biasio et al., 2011). However, this configuration places PCNA molecules loaded onto DNA in opposite directions, which is inconsistent with our current understanding of PCNA structure and function. Nevertheless, it is possible that two PCNA molecules could interact this way in *trans* in

the context of a replication factory. Determining whether two PCNA$^{SC}$(WT-79-79) molecules interact in this fashion could address whether PCNA trimers do in fact dimerize and whether dimerization is important for PCNA function during CRL4$^{Cdt2}$-mediated proteolysis. My preliminary data showed that having more than one IDCL on a single trimer is beneficial and that a single trimer can interact with other PCNA molecules. If confirmed, these data are consistent with a model in which PCNA dimerization becomes necessary for CRL4$^{Cdt2}$ function only when a separate IDCL is unavailable on the same PCNA trimer to which the substrate is bound, but this cooperation in *trans* between two PCNA trimers is not as efficient at supporting CRL4$^{Cdt2}$ function as cooperation between two subunits on the same trimer.

# Appendix A1

# Polη and the ATR checkpoint
# and
# Polη regulation by $CRL4^{Cdt2}$

The work presented in this appendix was partially conducted in Dr. W. Matthew Michael's lab at Harvard University, currently at The University of Southern California. Polη constructs and antibodies were generated in the Michael Lab, and the data presented in Figure A1.1 was performed in the Michael lab. Experiments presented in Figure A1.2 and Figure A1.3 were performed in the Walter lab.

## *A1.1 Introduction*

During the course of DNA replication, replicative polymerases can stall at sites of DNA damage. Unassisted, such obstacles lead to replisome dissociation at the site of damage, replication fork collapse, double-strand breaks, and ultimately to chromosome breaks and translocations. This genome instability can lead to cancer. Nonetheless, exposure to genotoxic mutagens is common to all cells. As a result, replication checkpoints have evolved to stabilize replication forks and halt the cell cycle, blocking mitosis until replication has resumed and preventing chromosome loss. Additionally, several specialized translesion synthesis (TLS) polymerases are able to bypass DNA lesions, allowing replication to continue in the face of damage and leaving the lesion behind for repair pathways to resolve. Together, replication checkpoints and TLS polymerases allow continued replication but prevent mitosis until replication is complete.

One mechanism of resuming replication at sites of stalled replicative polymerases, allowing S phase to resume, is TLS. The low processivity, low fidelity TLS polymerases lack 3'→5' exonuclease activity and have large active sites to accommodate DNA lesions. Polη specializes in bypassing cis-syn cyclobutane pyrimidine dimers (CPDs) (Johnson et al., 1999) and 6-4 photoproducts (Johnson et al., 2001) generated by UV damage, though it is also capable of bypassing other types of DNA lesions. Recently, Polη was also shown to have a role in PCNA ubiquitylation, which affects recruitment of

other TLS polymerases and homology-directed DNA repair (Durando et al., 2013). Despite its general low fidelity, Polη successfully bypasses double thymidine dimers by inserting two adenines opposite the adjoined bases (Johnson et al., 1999; McCulloch and Kunkel, 2008), earning it the distinction “error-free” when replicating this cognate lesion. A Polη deficiency causes Xeroderma pigmentosum variant (XPV) in humans (Masutani et al., 1999), a condition characterized by extreme susceptibility to skin cancer due to a high rate of mutagenesis following exposure to UV light.

In early *C. elegans* embryos, Polη is important for regulating developmental timing. Early embryos of many animals rely on a timing-specific sequence of events for proper development, so a delay in the cell cycle is detrimental to survival. Many organisms have evolved mechanisms to evade a checkpoint response during the early steps of development. In early *C. elegans* embryos, Polη is necessary for ATR checkpoint silencing (Holway et al., 2006). Additionally, in human XPV cells, which lack functional Polη, S phase arrest is lengthened and accompanied by increased checkpoint signaling (Bullock et al., 2001); adding Polη to these cells prevents delays in the cell cycle and excessive checkpoint signaling (Albertella, 2005). Together, these observations suggest a functional interaction between Polη and ATR checkpoint signaling.

Also in early *C. elegans* embryos, $CRL4^{Cdt2}$ targets Polη for destruction. (Kim and Michael, 2008). Polη is first SUMOylated, which protects it from proteolysis (Kim and Michael, 2008). Importantly, both deletion of the SUMOylation machinery and high amounts of MMS damage are necessary to detect Polη destruction (Kim and Michael, 2008). Polη is error-prone on undamaged DNA and must not be granted unlimited access to DNA during lesion bypass. Therefore, mechanisms must exist to ensure that a replicative polymerase replaces the TLS polymerase after bypass is complete. Some models propose that its low processivity causes the TLS polymerase to simply fall off the

DNA after bypass. Kim and Michael hypothesized that $CRL4^{Cdt2}$-dependent destruction of Polη might be necessary for polymerase exchange to the replicative polymerase after TLS. They hypothesized that Polη is protected from destruction briefly via SUMOylation on DNA, but it then becomes susceptible to $CRL4^{Cdt2}$-dependent destruction, which facilitates the polymerase switch (Kim and Michael, 2008). Sequence alignments show that *Xenopus* Polη contains elements of the conserved PIP box and basic +4 residue necessary for Cdt2-dependent destruction (Havens and Walter, 2009), suggesting that $CRL4^{Cdt2}$ regulation of Polη might be conserved in higher eukaryotes. In this appendix, I present my work studying the regulation of Polη by $CRL4^{Cdt2}$ and its interaction with ATR checkpoint signaling in *Xenopus* egg extract.

## *A1.2 Materials and Methods*

Extract preparation, replication assays, and chromatin spin downs are described above (Chapter 2, Chapter 3)

### A1.2.1 Plasmid construction, mutagenesis, and protein purification

PCNA purification is described in Chapter 3.

The construct for *Xenopus* Polη was previously donated to the Michael lab. From this construct, three overlapping His-tagged Polη fragments were produced for the generation of an antibody against *Xenopus* Polη. Recombinant *Xenopus* Polη was expressed and purified from *E. coli* according to published methods (Yagi et al., 2005).

### A1.2.2 Immunological Methods

Antibodies to PCNA, Cdt1, Rcc1, Orc2, phospho-Chk1, and Chk1 are described above (Chapter 2, Chapter 3). Anti-Polη antibodies were generating using the C-terminal fragments described above (A1.2.1) as an antigen for injection into two rabbits. The C-terminus of Polη shares the least sequence homology with other TLS polymerases. Polη

antibodies were produced by Cocalico Biologicals.

## *A1.3 Results*

### A1.3.1 Polη in crude S phase extract

To study Polη, both in relation to the S phase checkpoint and as a potential $CRL4^{Cdt2}$ substrate, I generated an antibody that works for Western blotting and immunodepletion in the Michael lab. I also generated recombinant full-length Polη (Yagi et al., 2005). Purified Polη had strand displacement activity immediately after purification, but any subsequent dialysis or other manipulations rendered it inactive, likely due to denaturation (data not shown). I therefore focused on studying endogenous Polη. In agreement with classification of UV damage as the cognate substrate of Polη, UV damage led to increased Polη binding sperm chromatin in crude S phase extract (Figure A1.1A). Additionally, Polη depletion from crude extract led to decreased replication of UV-damaged chromatin (Figure A1.1B, C), which we also expected since Polη is the TLS polymerase dedicated to bypassing UV-induced pyrimidine dimers. Due to our inability to purify an active recombinant Polη, I could not rescue the effect of the depletion.

### A1.3.2 Polη and ATR checkpoint signaling

We were particularly interested in the effects of Polη depletion or addition on ATR checkpoint signaling due to the relationship between Polη and checkpoint signaling in nematodes. However, depletion of Polη from extract did not affect checkpoint signaling (Figure A1.1D), and the addition of recombinant Polη to extract gave inconsistent results (data not shown), likely due to differences in recombinant Polη activity. Hence, I terminated this project.

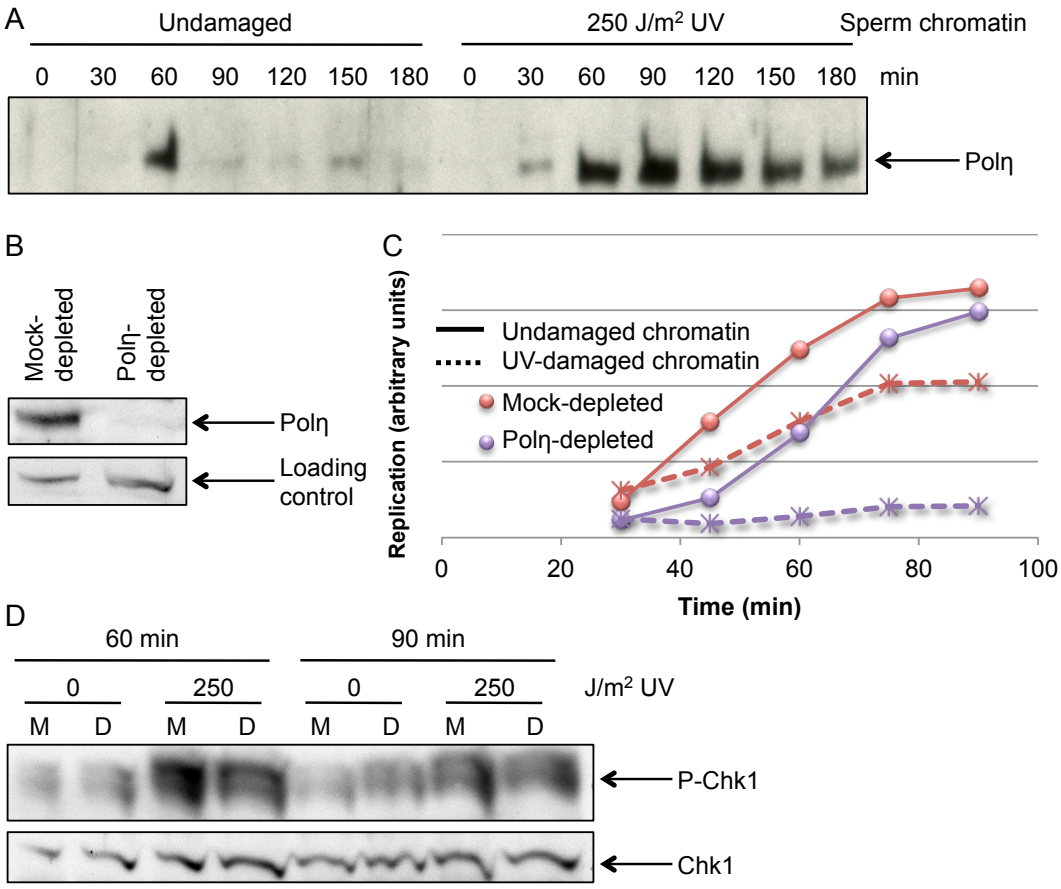



**Figure A1.1: Polη affects DNA replication of UV-damaged chromatin but not ATR checkpoint signaling. A)** Polη interacts strongly with replicating UV-damaged sperm chromatin and weakly with replicating undamaged sperm chromatin. **B)** A single-round Polη depletion removes most Polη from crude S phase extract. **C)** Depletion of Polη impairs DNA replication of sperm chromatin, especially UV-damaged sperm chromatin. **D)** Depletion of Polη does not affect ATR checkpoint signaling during replication of UV-damaged sperm chromatin as measured by Chk1 phosphorylation. M and D indicate mock-depleted and Polη-depleted, respectively. These experiments were performed in crude S phase extract in the Michael lab.

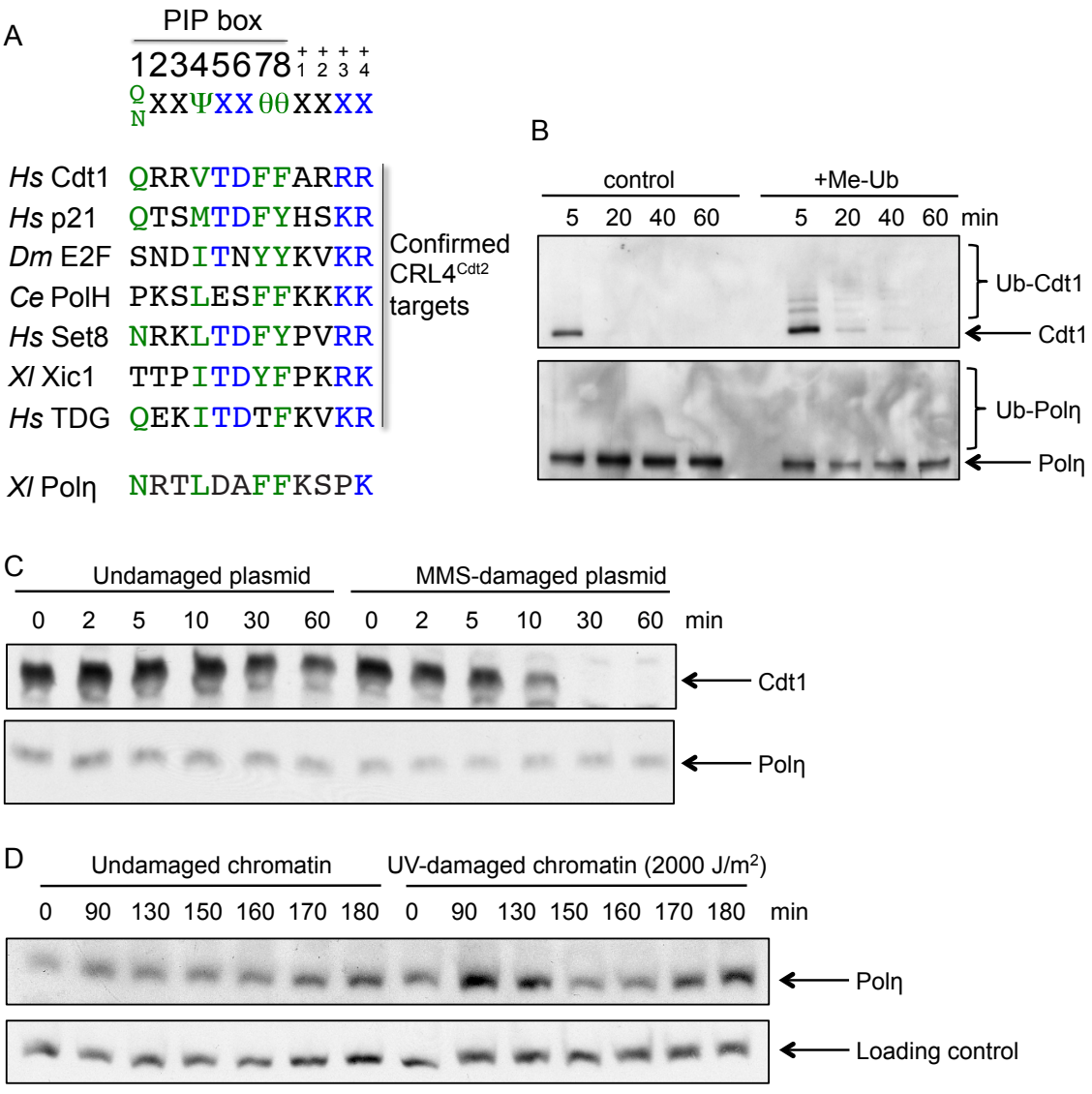
### A1.3.3 Putative CRL4$^{Cdt2}$ regulation of Polη

Because Polη is a CRL4$^{Cdt2}$ substrate in *C. elegans*, I set out to characterize its proteolysis in *Xenopus* egg extract and test the hypothesis that CRL4$^{Cdt2}$-dependent destruction is necessary to limit continued replication by Polη after lesion bypass. *Xenopus* Polη contains a putative PIP degron (Figure A1.2A); there is a basic residue four amino acids downstream of its previously established C-terminal PIP box (Acharya et al., 2008). I first looked for ubiquitylation of Polη on MMS-damaged linear DNA that I could easily recover from extract. This template supports the ubiquitylation of known CRL4$^{Cdt2}$ substrates Cdt1 (Chapter 2 and Figure A1.2B, upper panel) and TDG (Chapter 2). I optionally supplemented extract with methyl ubiquitin, which allows monoubiquitylation but prevents polyubiquitin chain formation and proteasome-dependent destruction. This template triggered Cdt1 destruction in the absence of methyl ubiquitin and resulted in Cdt1 ubiquitylation when methyl ubiquitin was included (Figure A1.2B, upper panel). However, Polη remained unmodified on chromatin (Figure A1.2B, lower panel). I also tested for destruction of Polη during DNA repair in HSS. I again used MMS-damaged DNA, which triggers proteolysis of other known substrates. However, Polη remained stable under conditions that stimulated rapid Cdt1 destruction (Figure A1.2). Finally, replication of both undamaged and UV-damaged sperm chromatin did not trigger any change in Polη levels (Figure A1.2C).

To independently confirm that Polη is not destroyed in a CRL4$^{Cdt2}$-dependent manner, I immunodepleted PCNA from HSS and replaced it with a PCNA mutant that supports substrate binding but not ligase recruitment (D122A/E124A) (Havens et al., 2012). I tested the effect of this mutant on levels of Polη associated with both undamaged and UV-damaged sperm chromatin and found no accumulation of Polη in either case (Figure A1.3). The ligase-recruitment dead version of PCNA led to increased Cdt1 binding under

**Figure A1.2: Polη is not ubiquitylated or destroyed under conditions that trigger Cdt1 ubiquitylation and destruction. A)** Sequence alignment of *Xenopus* Polη compared to known $CRL4^{Cdt2}$ substrates. **B)** Polη is not ubiquitylated on MMS-damaged linear DNA in HSS. Cdt1 is shown as a positive control for ubiquitylation. Methyl ubiquitin was optionally added to allow monoubiquitylation but prevent polyubiquitylation and destruction. **C)** An MMS-damaged plasmid does not trigger destruction of Polη in HSS. **D)** Polη is not destroyed during replication of UV-damaged or undamaged chromatin in LSS. These experiments were performed in the Walter lab.

**Figure A1.2** (continued)



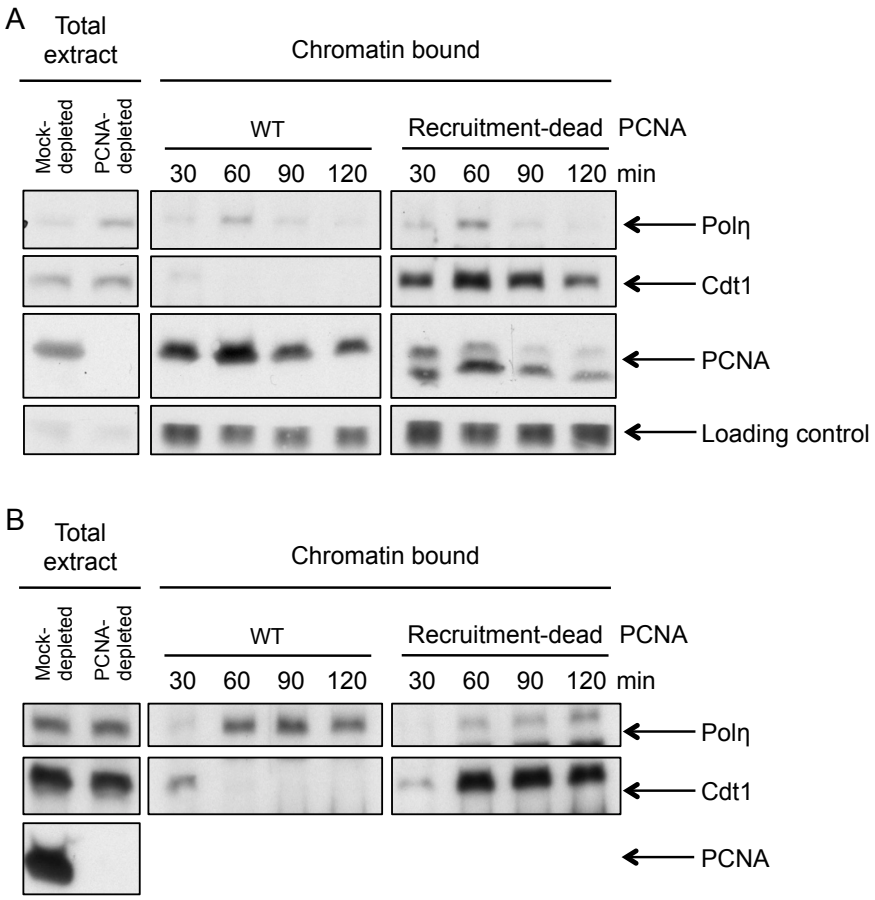



**Figure A1.3: Mutation of PCNA residues critical for CRL4$^{Cdt2}$ recruitment does not lead to the accumulation of Polη on chromatin. A)** Chromatin-associated proteins during DNA replication of normal sperm chromatin in LSS. Rcc1 is shown as a loading control. **B)** Chromatin-associated proteins during replication of UV-damaged sperm chromatin.

both conditions, as expected (Figure A1.3).

These new data indicate that Polη is not a target for the $CRL4^{Cdt2}$ ubiquitin ligase in *Xenopus* egg extract under conditions known to cause ubiquitylation and degradation of other $CRL4^{Cdt2}$ substrates, like Cdt1. Mass spectrometry results confirmed this final observation (Appendix A2). Though it remains a formal possibility that $CRL4^{Cdt2}$ somehow regulates Polη, it does not appear to do so following any of the types of damage tested.

The fact that Polη contains a conserved PIP box and basic +4 residue, but is apparently not a target for destruction in *Xenopus,* slightly expands our understanding of sequence requirements for $CRL4^{Cdt2}$ substrates. Within the PIP box, Polη contains an aspartate and alanine at positions five and six where most other substrates contain a TD motif (Figure A1.2). The PCNA-binding protein Fen1 contains two aspartates at these positions and does not contain a B+4 residue at all (Havens and Walter, 2009 and Chapter 3, Figure 3.1). Replacement of the first aspartate with threonine leads to tighter binding to PCNA. When a B+4 residue is also added, Fen1 becomes a substrate for $CRL4^{Cdt2}$ (Havens and Walter, 2009). Importantly, both changes are necessary to make Fen1 a target. Therefore, it is possible that PCNA-Polη binding is not strong enough to recruit Cdt2, so there is no destruction. Nonetheless, *C. elegans* Polη does not contain the important TD motif either, yet it is a target. It is thus possible that other residues compensate for the lack of a TD motif in nematode Polη that are not present in *Xenopus* Polη or that there is a difference in $CRL4^{Cdt2}$ requirements between species.

# Appendix A2

# A strategy for the proteomic identification of new $CRL4^{Cdt2}$ substrates in *Xenopus* egg extract

## *A2.1 Introduction*

Unlike other ubiquitin ligases, $CRL4^{Cdt2}$ functions specifically during DNA replication or following DNA damage. $CRL4^{Cdt2}$ negatively regulates the stability of many proteins whose presence in S phase is detrimental. The best-characterized substrate of $CRL4^{Cdt2}$ is Cdt1 (Jin et al., 2006), which functions to recruit the replicative MCM2-7 helicase during origin licensing. After cells progress into S phase, Cdt1 activity could facilitate detrimental origin licensing on regions of the genome that have already been replicated (Arias and Walter, 2005) so $CRL4^{Cdt2}$ targets Cdt1 for destruction at the start of S phase.

Fewer than ten $CRL4^{Cdt2}$ substrates are known, but all have a role in DNA replication or repair (Havens and Walter, 2011). Other substrates of $CRL4^{Cdt2}$ include *Drosophila* E2f1(Shibutani et al., 2008), human Set8 (Centore et al., 2010), *Xenopus* Xic1 (Kim et al., 2010), *S. pombe* Spd1 (Liu et al., 2003; Salguero et al., 2012), *C. elegans* POLH-1 (the DNA Polη homologue) (Kim et al., 2010) and p21 in multiple organisms (Kim et al., 2008). We recently characterized a new substrate, thymine DNA glycosylase (TDG), as well (Chapter 2). Destruction of each $CRL4^{Cdt2}$ substrate helps to ensure accurate DNA replication, and most $CRL4^{Cdt2}$ substrates are toxic during S phase. Addressing why each substrate must be destroyed has helped us understand important molecular details about DNA replication and repair. For example, the identification of Set8 as a substrate led to the model that H4K20 methylation during DNA replication somehow leads to relicensing (Beck et al., 2012). Before the discovery that Set8 is a $CRL4^{Cdt2}$ substrate, the relationship between H4K20 methylation and cell cycle progression was unknown. Therefore, I expect that identification of novel $CRL4^{Cdt2}$ substrates will similarly improve our knowledge of DNA replication and repair, as well as help us gain more insight to the activities of new substrates.

To identify new $CRL4^{Cdt2}$ substrates and expand our knowledge of DNA replication and repair, I used a novel screening method developed by Dr. Courtney Havens, Dr. Johannes Walter, and myself.

## *A2.2 Materials and Methods*

Extract preparation, immunological methods, DNA damage, tissue culture, and recovery of bead-bound DNA are described above (Chapter 2, Chapter 3). Recombinant PCNA was prepared as described in Chapter 3.

### A2.2.1 Preparation of samples for mass spectrometry

PCNA was immunodepleted from HSS and replaced with either wildtype or D122A/E124A PCNA. Bead-bound MMS-damaged DNA were added to the extract. At a time point that leads to near-complete degradation of Cdt1 on wildtype PCNA but minimal loss of Cdt1 on mutant PCNA, bead-bound DNA and associated proteins were isolated. Proteins were eluted in 1% SDS and then TCA precipitated: the eluate was adjusted to 20% TCA, vortexed, and incubated on ice for 20 minutes. Precipitated protein was recovered via centrifugation at max speed, 4°C, washed with cold 10% TCA, and then washed twice with cold acetone. The dried pellet was stored at -80°C and subsequently submitted to the Gygi lab for trypsinization, differential labeling and analysis.

### A2.2.2 Isolation of nuclear extract from HeLa cells

HeLa cells were cultured to confluency, harvested via trypsinization, and washed with PBS containing 1 mM $MgCl_2$. Cells were lysed on ice first in hypotonic buffer (1.5 mM $MgCl_2$, 10 mM KCl, 300 mM sucrose, 1 mM EDTA, 0.1% NP-40, 1 mM DTT, 1 mM $NaVO_4$, 10 mM HEPES, pH 7.9, and protease inhibitors) until they were permeable to the vital dye Trypan Blue. Cells were mixed via repeated pipetting to release the

cytoplasmic content and centrifuged to separate the cytoplasm from the insoluble fraction. The insoluble fraction containing the nuclei was permeabilized with hypertonic buffer (250 mM NaCl, 1 mM EDTA, 0.1% SDS, 0.5% sodium deoxycholate, 1% Triton X-100, 1 mM DTT, 1 mM $NaVO_4$, 1 mM NaF, 20 mM HEPES, pH 7.4, and protease inhibitors) on ice. The nuclei were mixed via repeated pipetting and nuclear contents were isolated via centrifugation at high speed. The nuclear lysate was in the supernatant of this final spin.

## *A2.3 Results and Troubleshooting*

### A2.3.1 Experimental set up

To identify new $CRL4^{Cdt2}$ substrates, I used a mutant of PCNA that contains mutations in the part of the protein that is predicted to make contacts with Cdt2 (D122A/E124A, dotted in Figure 1.2B). The two mutated acidic residues normally sandwich the B+4 residue on the substrate's PIP degron. This mutant does not affect substrate binding, but it prevents ligase recruitment, thereby stabilizing $CRL4^{Cdt2}$ substrates on DNA (Havens et al., 2012). After completing this screen, Dr. Courtney Havens found that mutation of D122A alone leads to the same effect (Havens et al., 2012). Substrates like Cdt1 and human Set8 rapidly bind a short, MMS-damaged DNA fragment (1kb). They then become modified with ubiquitin and are subsequently degraded (Centore et al., 2010; Havens and Walter, 2009). Ubiquitylation and degradation depend on proper $CRL4^{Cdt2}$ recruitment. Since the D122A/E124A mutant of PCNA fails to make the necessary contacts with Cdt2, substrates like Cdt1 remain bound to DNA in samples containing the mutant protein. Thus, comparing DNA-bound proteins in the presence of wildtype and D122A/E124A PCNA should lead to the identification of other $CRL4^{Cdt2}$ targets, as they will be enriched on DNA in the presence of the mutant PCNA.

I immunodepleted PCNA from HSS and replaced it with the D122A/E124A mutant and immobilized MMS-treated DNA. I recovered the DNA and eluted the samples in SDS buffer before TCA precipitating them and submitting them for differential labeling with a heavy and light isotope followed by mass spectrometry.

### A2.3.2 Optimization

To optimize the amount of $CRL4^{Cdt2}$ substrates present in the samples relative to other proteins, I varied the salt and detergent conditions during sample isolation. I used Cdt1 as an example substrate and planned all optimization around the amount of Cdt1 relative to background proteins in each sample. At the lowest salt concentration tested (50mM KCl), Cdt1 peptide hits were low compared to other DNA binding proteins and to some soluble proteins as well (Table 1). I made various adjustments to the experiment to enrich for substrates over abundant background proteins. Increasing salt concentrations during the chromatin spin downs reduced background levels at the cost of decreased DNA-associated Cdt1 (Figure A2.1A, B, showing both WT PCNA and D122A/E124A, labeled AA), but did not have a large affect on Cdt1 peptide hits or identify any additional Cdt2 targets (Table 1 and data not shown). With more stringent isolation conditions (100 mM and 150 mM KCl), the single-stranded DNA binding protein RPA bound with much greater efficiency than a background, cytoplasmic protein (β-actin) suggesting that the additional salt present was successful in eliminating nonspecific binding (Table 1).

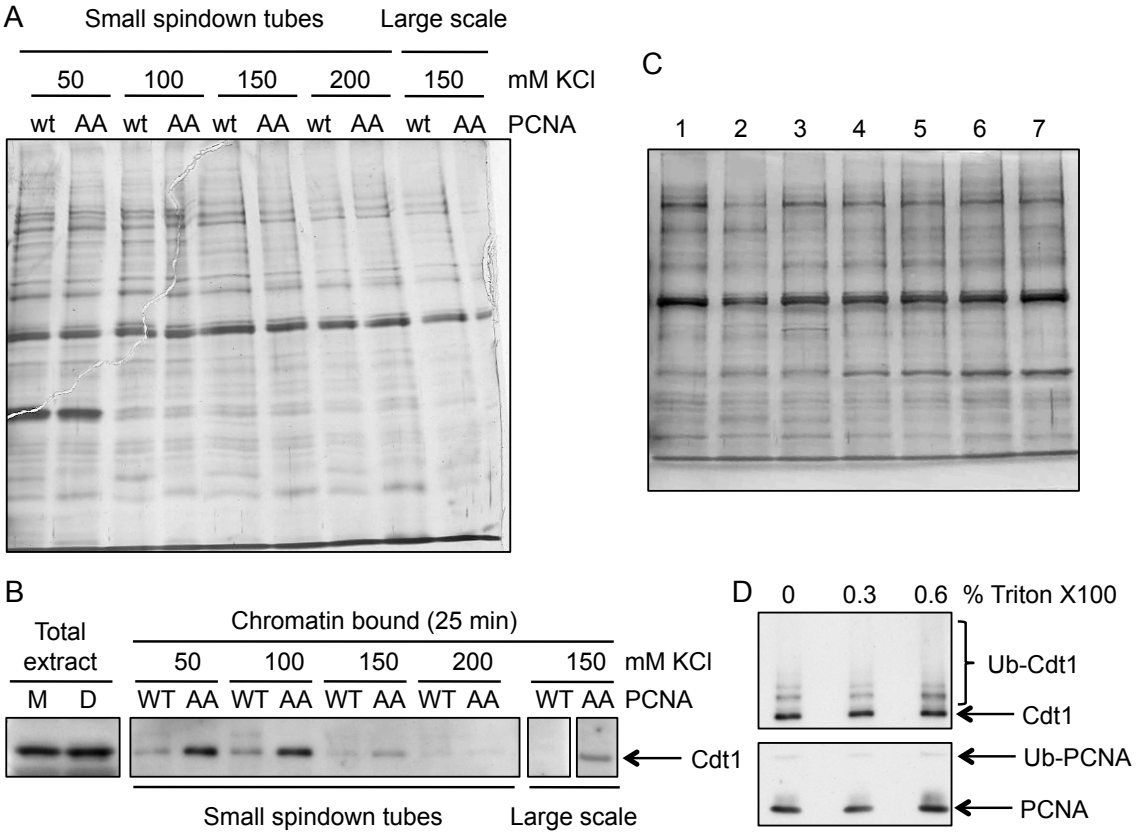
I also tried adding undamaged competitor DNA to the reactions to titrate away some non-specific chromatin binding proteins from the damaged immobilized DNA. The mass spectrometer can only detect a finite number of peptides. If the sample is full of DNA-binding proteins that bind both damaged and undamaged DNA, they will overwhelm the number of specific hits. However, addition of undamaged competitor DNA did not affect total protein in the recovered DNA fraction as measured by silver staining

**Table 1: Total peptide hits with increasing salt concentrations**

| Salt concentration: | 50 | 100 | 150 | mM KCl |
|---|---|---|---|---|
| Cdt2 | 10 | 10 | 3 | |
| Cdt1 | 1 | 4 | 0 | |
| RPA | 354 | 333 | 570 | |
| β-actin | 466 | 84 | 171 | |
| | | | | |
| DNA-specificity | 0.76 | 3.96 | 3.33 | |

**Figure A2.1: Optimization of sample preparation for differential labeling and mass spectrometry. A)** The effect of increasing salt concentration (KCl) on total DNA-associated protein as detected by silver staining. Here, WT indicates wildtype PCNA, and AA indicates ligase recruitment-defective PCNA (D122A/E124A). **B)** The effect of increasing salt concentration (KCl) on DNA-associated Cdt1. This Western blot corresponds to (A). **C)** Undamaged competitor plasmid (2 μg) was added to titrate away non-specific DNA binding proteins. Lanes are as follows: 1) Buffer was added to DNA-bound beads before adding HSS to start the reaction, 2) Plasmid was mixed with DNA-bound beads before adding HSS to start the reaction, 3) Plasmid was added to the entire reaction one minute before isolating bead-bound DNA, 4) Plasmid was added to the diluted reaction before separating through the sucrose cushion, 5) Buffer was added to the wash solution after separating the beads through the sucrose cushion, 6) Plasmid was added to the wash solution after separating the beads through the sucrose cushion, 7) No plasmid or buffer was included. **D)** The effect of increasing Triton X-100 on Cdt1 and PCNA recovery. This experiment was performed with 150 mM KCl.

**Figure A2.1** (continued)



(Figure A2.1C), so I did not pursue this approach with our samples submitted for mass spectrometry. Finally, I optimized the amount of Triton X-100 (Figure A2.1D).

### A2.3.3 Mass spectrometry results

Despite the low number of total Cdt1 peptides identified during mass spectrometry (Table 1), the difference between Cdt1 in samples containing wildtype PCNA compared to the ligase recruitment-dead PCNA (D122A/E124A) was measurable (four hits in samples containing WT PCNA and no hits in samples containing the ligase recruitment-dead PCNA). The final differential labeling results indicated that Cdt1 was enriched 15.5-fold in samples containing D122A/E124A PCNA, as expected (Table 2). Cdt2 and Ddb1, the $CRL4^{Cdt2}$ adapter protein, drastically decreased in samples containing the mutant PCNA, as well (Table 2 for Cdt2 comparison and number of peptide hits only—27 in the wildtype samples and 0 in the mutant samples—for Ddb1), indicating that both ligase recruitment and substrate destruction were defective.

Of the other proteins enriched on DNA in samples containing mutant PCNA, a few proteins stood out due to their roles in DNA replication or repair: Sub1/PC4, HMG1, and YB-1 (in bold font in Table 2). I identified a putative PIP box in Sub1/PC4, which has been shown to be recruited to sites of DNA damage (Mortusewicz et al., 2008) (Figure A2.2A). Though there is no B+4 residue following the PIP box of Sub1/PC4, there are several basic residues upstream of the PIP box. Upstream basic residues play a role in $CRL4^{Cdt2}$ function (Havens et al., 2012), so it is plausible that they could circumvent the requirement for downstream basic residues for some substrates. I generated a construct containing the Sub1/PC4 gene under control of the Sp6 promoter to generate *in vitro* transcribed and translated Sub1/PC4. I added this Sub1/PC4 to HSS containing an MMS-damaged plasmid to trigger destruction of $CRL4^{Cdt2}$ substrates. *In vitro* transcribed and translated Sub1/PC4 remained stable over the course of an hour in extract

**Table 2: Proteins enriched in samples containing CRL4$^{Cdt2}$ recruitment-dead PCNA or containing wildtype PCNA. Note:** The PCNA ratio does not include peptides that contain the mutation (6 peptides out of 73 total peptides in the sample containing mutant PCNA). The WT sample contains 63 total peptide hits.

| Protein | $Log_2$ Ratio | Fold Enrichment |
|---|---|---|
| Higher in AA samples | | |
| **YB-2** | 1.3 | 2.5 |
| **YB-1** | 1.3 | 2.5 |
| MCM4B | 1.4 | 2.7 |
| zMCM6A | 1.5 | 2.7 |
| glutaminyl-tRNA synthetase | 1.5 | 2.7 |
| MCM5A | 1.5 | 2.8 |
| mMCM6 | 1.5 | 2.8 |
| MCM2 | 1.5 | 2.9 |
| MCM3 | 1.5 | 2.9 |
| MCM5B | 1.6 | 2.9 |
| MCM7A | 1.6 | 3.0 |
| methionyl-tRNA synthetase | 1.7 | 3.2 |
| importin alpha 1 | 1.8 | 3.4 |
| MCM7B | 1.8 | 3.4 |
| glutathione S-transferase mu 2 | 2.0 | 3.9 |
| **SUB1/PC4** | 2.0 | 4.1 |
| **HMG1** | 2.8 | 7.0 |
| Cdt1 | 4.0 | 15.5 |
| No difference | | |
| PCNA | -0.2 | 0.8 |
| Polymerase delta 1 | -0.2 | 0.9 |
| Polymerase delta 2 | -0.1 | 0.9 |
| Polymerase delta 3 | -0.2 | 0.8 |
| Higher in WT samples (inverse enrichment reported) | | |
| Cdt2 | -3.5 | 11.5 |
| DNA polymerase epsilon | -2.7 | 6.6 |
| DNA polymerase epsilon 2 | -2.2 | 4.6 |
| Histone H4 | -2.0 | 4.0 |
| Orc5 | -1.9 | 3.8 |
| DNA ligase 1 | -1.7 | 3.1 |
| RNA binding motif protein | -1.5 | 2.9 |
| MCM10 | -1.2 | 2.4 |

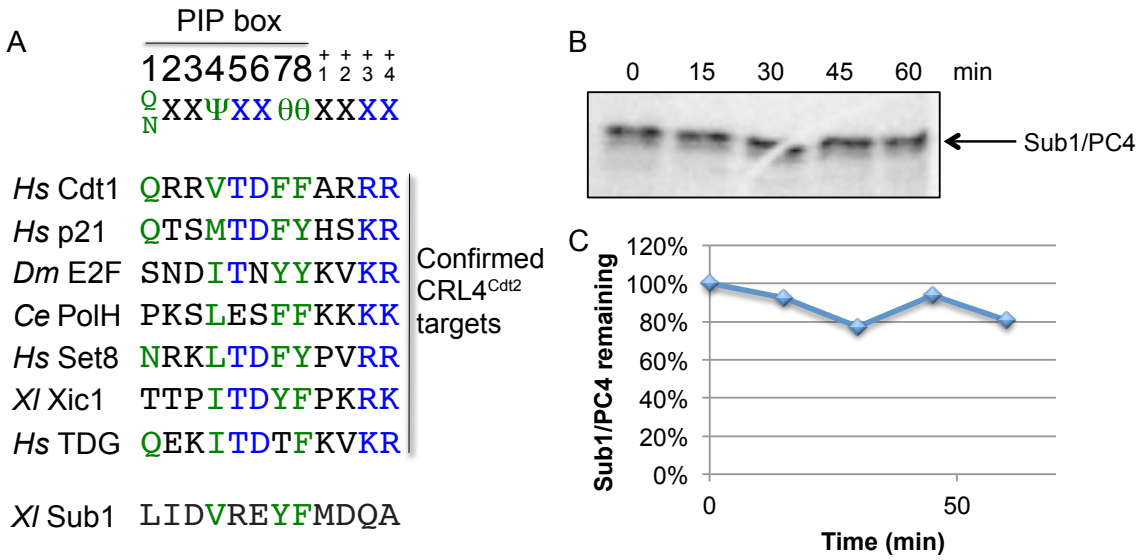



**Figure A2.2: In vitro transcribed and translated Sub1/PC4 is not destroyed during DNA repair in HSS. A)** The putative PIP box of Sub1/PC4 is shown aligned to confirmed $CRL4^{Cdt2}$ substrates. **B)** Sub1/PC4 levels do not change in HSS containing MMS-damaged plasmid. **C)** Quantification of (B).

(Figure A2.2B, C), however, suggesting that it is not destroyed in a DNA damage-dependent fashion. It remains possible that the *in vitro* transcribed and translated form of Sub1/PC4 folds improperly or that its removal from DNA was impaired without affecting the overall pool of protein, but it seems unlikely that Sub1/PC4 is a $CRL4^{Cdt2}$ substrate based on these data alone.

HMG1 and YB-1 are also known to bind chromatin and were enriched in samples containing D122A/E124A PCNA. However, another group looked for PCNA binding of HMG1 and found no association (Ise et al., 1999), and though YB-1 has been shown to interact with PCNA (Ise et al., 1999), I could not identify a PIP box and therefore did not pursue these potential targets.

Despite the inability of the mass spectrometry results to identify new $CRL4^{Cdt2}$ targets, it did identify several DNA-binding proteins that were greatly enriched in the samples containing wildtype PCNA (Table 2). The most likely explanation for this observation is that a variety of $CRL4^{Cdt2}$ substrates that were bound to the ligase recruitment-defective PCNA prevented other PIP box-containing proteins from binding DNA through PCNA. Among these proteins enriched in the wildtype sample were several origin recognition complex (ORC) subunits and the replicative DNA polymerase epsilon (Pol$\varepsilon$) (Table 2). Notably, there was no difference in Pol$\delta$ binding. MCM7, on the other hand, increased appreciably in the presence of the mutant PCNA (Table 2). It is possible that Cdt1, which is stabilized on DNA in the presence of the mutant PCNA, triggers additional recruitment of the MCM helicase. I confirmed that the PCNA depletion at the start of the experiment was successful and that these DNA replication proteins followed the trends indicated by mass spectrometry by Western blotting (Figure A2.3).

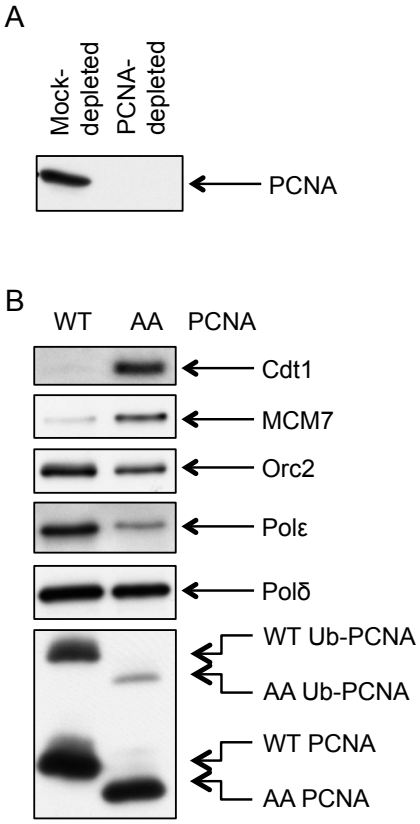



**Figure A2.3: Confirmation of mass spectrometry results. A)** The PCNA depletion in samples analyzed via mass spectrometry was complete. **B)** A comparison of replication proteins in samples containing wildtype PCNA or $CRL4^{Cdt2}$-recruitment dead PCNA (labeled AA to indicate the D122A/E124A ligase recruitment-defective mutant). Notably, the PCNA antibody does not detect the mutant PCNA (AA) as well as wildtype PCNA, though samples containing mutant PCNA contain equal or greater amounts of total PCNA (Table 2).

### A2.3.4 Supplementation of the screening method with human cell lysates

To advance this screening method further, I attempted to supplement *Xenopus* egg extract with nuclear lysate from human cells. Many proteins are not expressed during the early cleavage divisions of *Xenopus* development that our extracts mimic. Transcription of most zygotic genes does not begin until the mid-blastula transition after the first 12 cleavage divisions, and controlled regulation of gene expression begins during gastrulation (Bouvet et al., 1994; Rupp and Weintraub, 1991; Veenstra et al., 1999; Wormington and Brown, 1983). We reasoned that some developmentally regulated genes might not be expressed in our extract, so they would not appear in our screen. $CRL4^{Cdt2}$ could be regulated developmentally as well, allowing it to target different substrates for destruction in different environments. Thus, I generated nuclear lysate from HeLa cells (Singh et al., 1998), which contains substrates for $CRL4^{Cdt2}$-dependent destruction that might not be present in *Xenopus* egg extract. Additionally, factors in HeLa cells might regulate $CRL4^{Cdt2}$ differently. All elements of the $CRL4^{Cdt2}$ pathway are interchangeable (Centore et al., 2010; Havens and Walter, 2009), so the *Xenopus* machinery should destroy human substrates. I successfully isolated nuclear extract (Figure A2.4A). However, the nuclear extract impaired $CRL4^{Cdt2}$ in extracts containing nuclear lysate, likely due to the detergent present in lysis buffer. The addition of nuclear lysates to HSS inhibited destruction of endogenous *Xenopus* Cdt1, human Cdt1, and human Set8 (Figure A2.4), even though recombinant human Cdt1 and recombinant human Set8 can be destroyed in HSS (Centore et al., 2010; Havens and Walter, 2009).

## *A2.4 Future Directions*

I designed a new screening method to identify $CRL4^{Cdt2}$ substrates in *Xenopus* egg extract based on the differential retention of $CRL4^{Cdt2}$ substrates on DNA in the presence of a ligase recruitment-defective version of PCNA. Although I did not identify any new

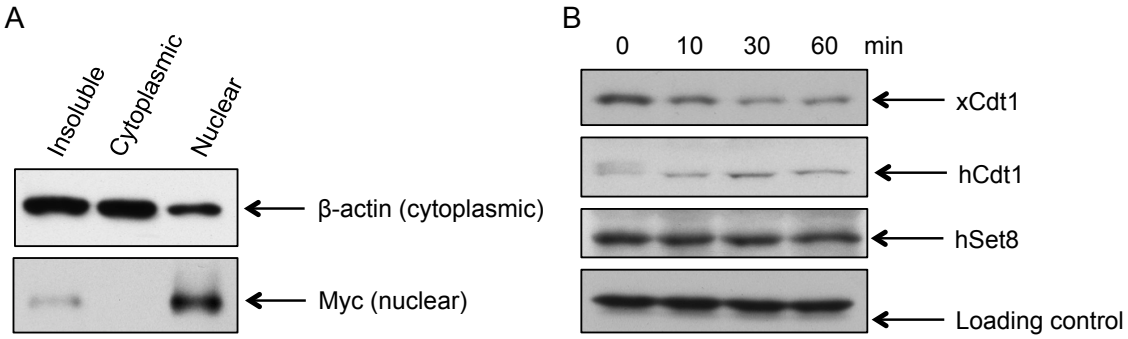



**Figure A2.4: Attempts to identify new CRL4$^{Cdt2}$ substrates from human cells using the *Xenopus* egg extract system. A)** Isolation of nuclear lysate. **B)** Addition of nuclear lysate (25% by volume) to HSS eliminates CRL4$^{Cdt2}$ function.

substrates, it is possible that this screening technique can be modified to identify new $CRL4^{Cdt2}$ substrates. Firstly, not all $CRL4^{Cdt2}$ substrates are present in extracts at appreciable levels. For example, TDG is barely detectable in extracts or early embryos (Chapter 2). Different substrates are also destroyed at different rates; TDG is destroyed much more slowly than Cdt1 (Chapter 2). This screen, as conducted, did not identify DNA-bound TDG in samples containing wildtype or D122A/E124A. Repeating the experiment at later time points could lead to a different pool of DNA-bound proteins to compare, leading to the identification of imperfect substrates like TDG (Chapter 2) if they are abundant enough in extract to allow detection by mass spectrometry. Additionally, supplementation of extracts with either human cell lysates or lysates from developing *Xenopus* embryos could lead to the identification of novel substrates if we can do so without impairing the ability of the extract to trigger destruction. A detergent-free protocol for isolating nuclear extracts has been used successfully as an extract system for transcription and might be usable in this scheme as well (Carey et al., 2009; Dignam et al., 1983). Finally, using the HSS/NPE system instead of HSS alone could lead to more robust destruction of $CRL4^{Cdt2}$ substrates and cope with dilution of the extract machinery better. In the case of TDG, destruction is barely detectable until about an hour after the addition of MMS-damaged plasmid in HSS, but there is a clear difference in protein levels by 15 minutes in HSS/NPE (Chapter 2).

It might also be of interest to determine why Polε levels on chromatin are so much lower in samples containing D122A/E124A PCNA. One hypothesis is that competition for PCNA binding by the nondegraded $CRL4^{Cdt2}$ substrates causes decreased levels of Polε. This explanation is not in complete agreement with the literature, however, as Polε binds not only to the IDCL of PCNA but also to another region of the clamp, and the IDCL is more important for Polδ binding (Eissenberg et al., 1997). The PCNA mutant did not

affect Polδ levels on DNA (Table 2). One simple approach to address this hypothesis is to titrate a p21 PIP box peptide into extract. This peptide will bind the IDCL of PCNA, prohibiting binding of other PIP box containing proteins, such as Polε and Polδ. If my hypothesis is correct, I expect that the PIP box peptide will block Polε binding at a much lower concentration than is necessary to block Polδ binding. This would explain the results of the spin down experiments: Polδ might simply bind to PCNA with a higher affinity than Polε and would therefore be more competitive despite the presence of nondegraded $CRL4^{Cdt2}$ substrates. Preliminary data from this experiment was inconclusive (data not shown). However, this hypothesis might prove incorrect, and then identifying an alternative explanation for the extreme difference in Polε levels between the wildtype and D122A/E124A PCNA samples, despite similar Polδ levels, could point toward unknown regulatory mechanisms during DNA replication and repair.

# Appendix A3

# Additional studies on thymine DNA glycosylase

This appendix provides additional material that corresponds to Chapter 2. The motivation for these experiments is outlined in Chapter 2, Section 2.1.

## *A3.1 Materials and Methods*

All methods used in this section were identical to those used in Chapter 2 except those described below.

### A3.1.1 Plasmid construction, mutagenesis, and protein purification

Constructs encoding human TDG were provided by Dr. Primo Schär from the University of Basel in Switzerland. The human TDG gene was transferred to a pIRES2-EGFP vector through PCR with the primers 5'-AAAAAGCTAGCCACCATGGATTACAAGGAT GACGATGACAAGGATTACAAGGATGACGATGACAAGGAAGCGGAGAACGCGGGCA GC-3' and 5'-TTTTTGGAATTCTAAG CATGGCTTTCTTCTTCC-3' and an NheI (NEB) and EcoRI (NEB) restriction enzyme digest. The final construct encodes Flag-tagged hTDG and GFP on two separate transcripts. To insert human TDG into a vector that does not co-express GFP, the human TDG gene was mistakenly inserted into a vector thought to be pIRES2-EGFP using the same primers and restriction enzymes. Importantly, the TDG coding region of this final plasmid was identical to that of pIRES2-hTDG-EGFP described above.

All point mutants described here (human TDG ΔPIP, *Xenopus* TDG ΔDBD, and *Xenopus* TDG ΔSUMO/ΔDBD) were generated with a QuikChange II Site-Directed Mutagenesis Kit (Agilent).

### A3.1.2 Human cell tissue culture and transfection

T98G cells were maintained in MEM (Invitrogen). Cells were transfected using the Effectene transfection reagent (Qiagen) or the Lipofectamine transfection reagent (Invitrogen) according to the manufacturers' protocols. The amount of DNA, transfection

reagent, and additional buffers were varied according to the manufacturers 'protocols to optimize protein expression. However, this optimization needs to be repeated and fine-tuned for future transfections.

For cell cycle analysis, cells were fixed in 4% paraformaldehyde for 10 minutes at room temperature. Cells were then permeabilized in PBS with 0.1% Triton for five minutes on ice. Triton was removed through a wash in PBS, and cells were resuspended in PBS containing 15 µg/mL propidium iodide and 200 µg/mL RNase. DNA content was analyzed with a FACS machine (BD). Fixation with ethanol could not be used in these experiments because the GFP signal would be quenched, and GFP would leak out of cells during fixation/permeabilization.

## *A3.2 Results and Discussion*

### A3.2.1 How the DNA-binding domain of TDG contributes to proteolysis

Because mutation of the TDG PIP box does not affect DNA binding (Chapter 2, Figure 2.11), and because the PIP box of TDG is weaker than most $CRL4^{Cdt2}$ substrates (Chapter 2), we hypothesized that the DNA binding domain might be cooperate with the PIP box for docking onto chromatin. If this hypothesis is correct, then abolishing the initial interaction between TDG and DNA by mutation of the DNA binding domain should stabilize the protein, as TDG would not be able to form a productive interaction with PCNA or $CRL4^{Cdt2}$ in the absence of DNA binding. To test this hypothesis, we attempted to generate a TDG mutant that cannot associate with DNA. Based on the crystal structure of TDG associated with DNA (Figure 1.4), we mutated three residues within the catalytic core of TDG that interact directly with DNA (Maiti et al., 2008) to generate TDG ΔDBD (ΔDNA-binding domain; K265A/K279A/K281A), both in the context of a wildtype TDG and TDG ΔSUMO.

The base hydrolysis activity of TDG ΔDBD on a uracil opposite guanine decreased to approximately 1/3 that of wildtype TDG (Figure A3.1A, B). Increasing the concentration of TDG ΔDBD in the activity assay restored most activity, suggesting that catalytic activity is not abolished in this mutant (Figure A3.1A, B). The association between MMS-damaged DNA and TDG ΔSUMO/ΔDBD was weakened, as expected (Figure A3.1C). Nonetheless, destruction of TDG ΔDBD occurred normally (Figure A3.1D, compare lanes 1-4 to lanes 5-8), suggesting that either 1) the DNA binding domain mutation does not eliminate the initial TDG binding step and instead only affects the stability of the DNA-TDG association, or 2) an initial association between TDG and DNA, independently of the association between TDG and PCNA, is not required for destruction. I also found that TDG destruction does not require the catalytic activity of TDG (Figure A3.1D, lanes 9-12), but it is possible that catalytic activity is required in another context, such as during TDG destruction in S phase.

We chose our DNA-binding domain mutant based on the published crystal structure of the TDG catalytic core in a complex with DNA. However, the N-terminus of TDG is involved in DNA binding. Therefore, it is likely that residues in TDG's N-terminus contribute to the TDG-DNA association as well. Since the exact DNA binding domain of TDG has not been characterized, and there is no available crystal structure of the N-terminus, we cannot say with confidence that the direct association between TDG and DNA is not required for TDG proteolysis. It is possible that a partial N-terminal truncation mutant in combination with the ΔDBD mutation used here would eliminate any additional TDG binding to DNA. However, the PIP degron is in the N-terminal domain, so even a partial N-terminal truncation could affect proteolysis directly. A full mutational analysis of the TDG N-terminus would be necessary to fully test our hypothesis regarding TDG

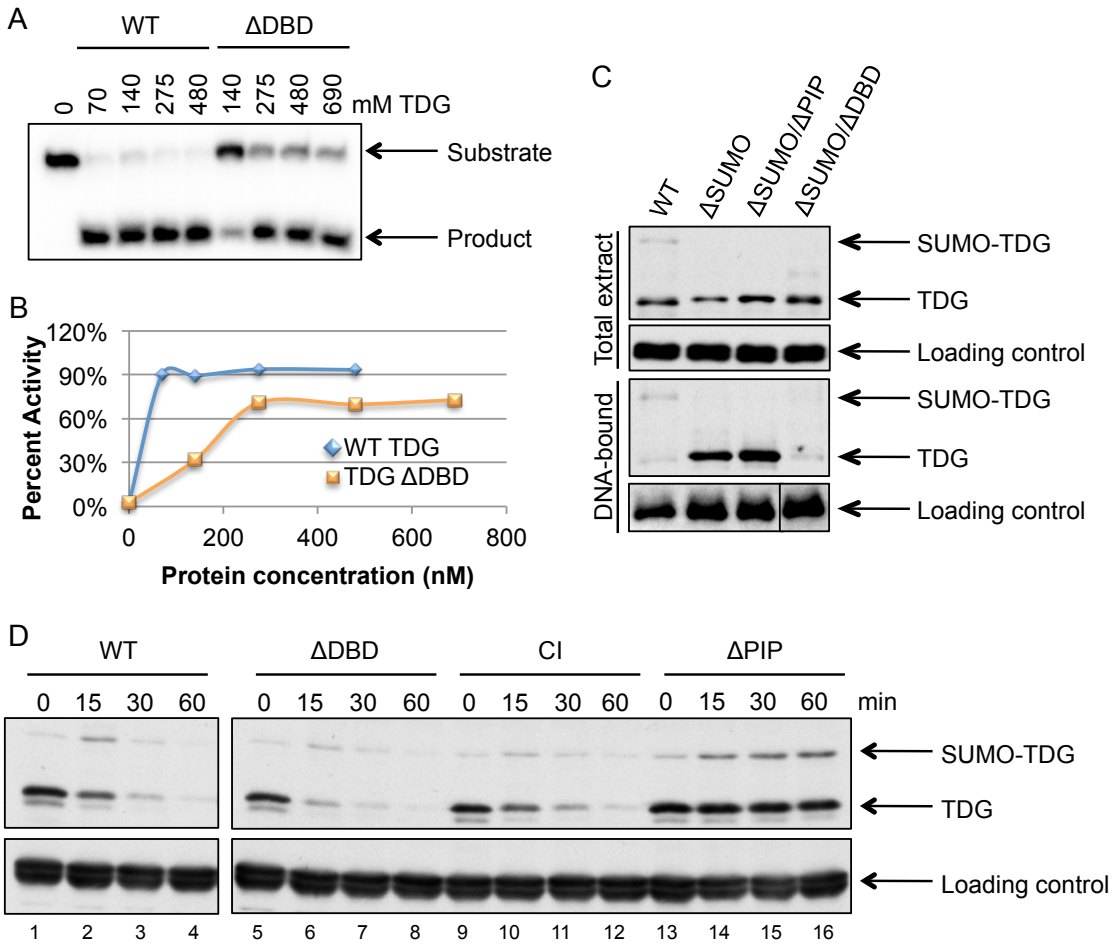



**Figure A3.1: Mutation of the DNA binding domain of TDG does not affect CRL4$^{Cdt2}$-dependent proteolysis. A)** TDG ΔDBD is only ~1/3 as active in a base hydrolysis assay as WT TDG at a concentration that results in full activity of the wildtype protein. Increasing the concentration of TDG ΔDBD partially rescues this defect, indicating the active site of the protein is not impaired. **B)** Quantification of (A). **C)** TDG ΔSUMO/ΔDBD does not associate with DNA as well as TDG ΔSUMO or TDG ΔSUMO/ΔPIP, which associate similarly. Lanes 1 and 2 (wildtype and ΔSUMO) were previously shown in Figure 2.5. A lane was skipped on the Orc2 blot shown as a loading control. **D)** TDG ΔDBD and the catalytic inactive TDG are destroyed normally during DNA repair.

binding to DNA.

### A3.2.2 The effects of expressing non-degradable TDG in human cells

High overexpression of TDG (20-30-fold) causes S phase arrest in 293T cells, and such high overexpression blocks cell proliferation (Hardeland et al., 2007). We therefore suspected that specifically disrupting the ability of cells to destroy TDG during S phase might have a similar effect. I transfected T98G cells with Flag-hTDG or the PIP box mutant. Transfection led to the expression of at least two TDG isoforms in T98G cells (Figure A3.2A). One of the two isoforms lost the N-terminal Flag tag and could be a cleavage product of the higher migrating isoform (Figure A3.2A). Transfection with Flag hTDG ΔPIP led to consistently higher protein expression than transfection with wildtype Flag hTDG despite transfection with the same plasmid concentrations and transfection reagents (Figure A3.2A). This is consistent with a failure to destroy Flag hTDG ΔPIP at stages of the cell cycle where endogenous TDG and wildtype Flag hTDG are normally destroyed.

Additionally, I generated another construct in which the open reading frame for hTDG also contains an internal ribosomal entry site followed by the GFP gene, so all transfected cells also express GFP, allowing us to measure transfection efficiency and look at the cell cycle profiles of transfected cells only. It also allows better comparison of transfection efficiency because GFP is not regulated like TDG, so we can look at total GFP intensity as a measure of transfection. Transfection with this TDG construct was successful. Approximately 60% of cells were transfected in each sample based on observation of GFP-positive cells. Lysates of cells transfected with pIRES2-hTDG-EGFP contained Flag-tagged TDG, and Flag-hTDG ΔPIP accumulated to higher levels than the wildtype protein despite equal transfection as measured by GFP intensity (data not shown, GFP measured by observation only).

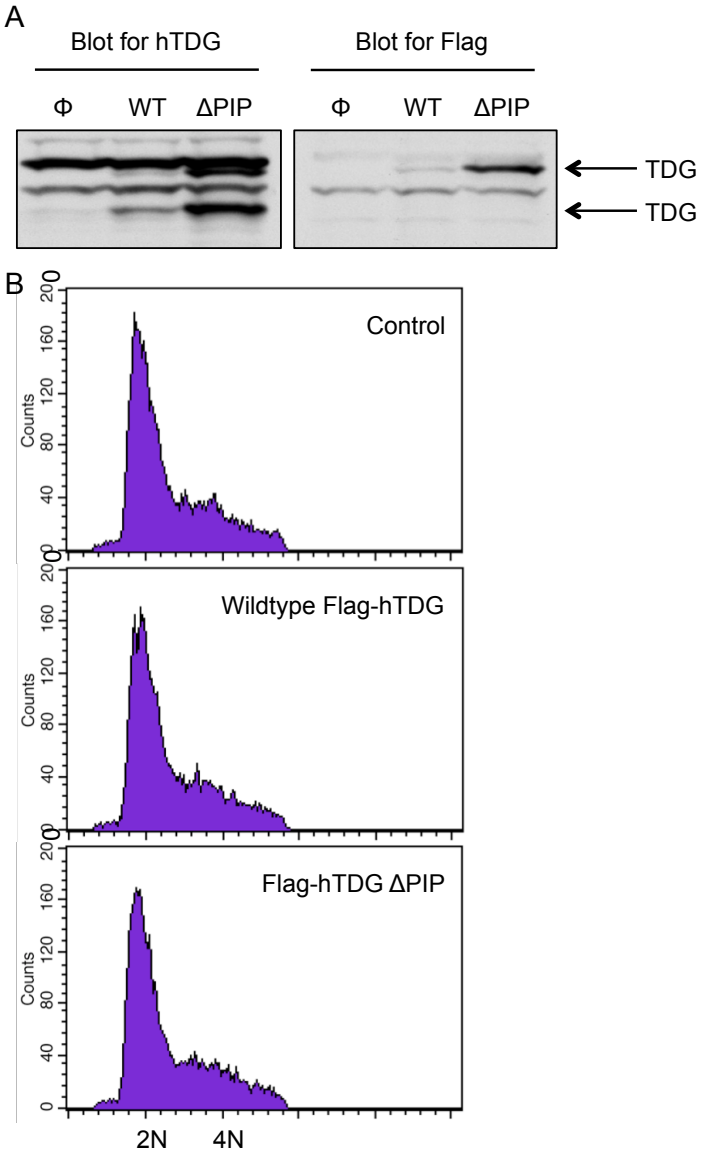



**Figure A3.2: Transfection of T98G cells with Flag-hTDG. A)** Flag-hTDG is expressed in T98G cells, and Flag-hTDG ΔPIP accumulates to even higher levels. Φ indicates transfection with an empty vector. When GFP was included in the construct, control cells were GFP-positive but did not express TDG. **B)** The cell cycle profiles of transfected cells are shown. Though Flag-hTDG ΔPIP does not appear to affect the cell cycle, these cell cycle profiles display no clear G2 peak or S phase.

Preliminary observation of the cell cycle profile after transfection with the construct that contains both hTDG and GFP was inconclusive (Figure A3.2B). The cell cycle profiles do not have a clear G2 peak, and I was unable to select for only GFP-expressing cells during FACS analysis. Cell cycle profiles in cells expressing the non-degradable hTDG mutant (Flag hTDG ΔPIP) are nearly identical to those of cells expressing wildtype Flag hTDG; I did not observe any accumulation in S phase after Flag-hTDG had been expressed for 2-3 cell cycles (Figure A3.2B). However, the experimental conditions need to be improved to allow for cleaner cell cycle profiles and clear detection of GFP during FACS.

### A3.2.3 Human TDG in *Xenopus* egg extract

Before generating *Xenopus* TDG, I added recombinant human TDG to *Xenopus* egg extract and monitored destruction. Under conditions where I have since seen destruction of *Xenopus* TDG, human TDG remained stable, though it was quickly SUMOylated like the *Xenopus* protein (Figure A3.3A). However, when I added a surplus of human Cdt2, I saw a modest decrease in hTDG ΔSUMO levels in response to damaged DNA (Figure A3.3B). *Xenopus* TDG destruction is much slower in HSS than in HSS/NPE, but *Xenopus* TDG is still destroyed in a PCNA-dependent manner without the addition of Cdt2 in HSS (Figure A3.3C). Other known human $CRL4^{Cdt2}$ substrates, like hCdt1 and hSet8, are destroyed in HSS despite the species difference (Centore et al., 2010; Havens and Walter, 2009). This means that there is some difference between TDG destruction and destruction of other known substrates. It is possible that species-specific factors regulate $CRL4^{Cdt2}$-dependent proteolysis.

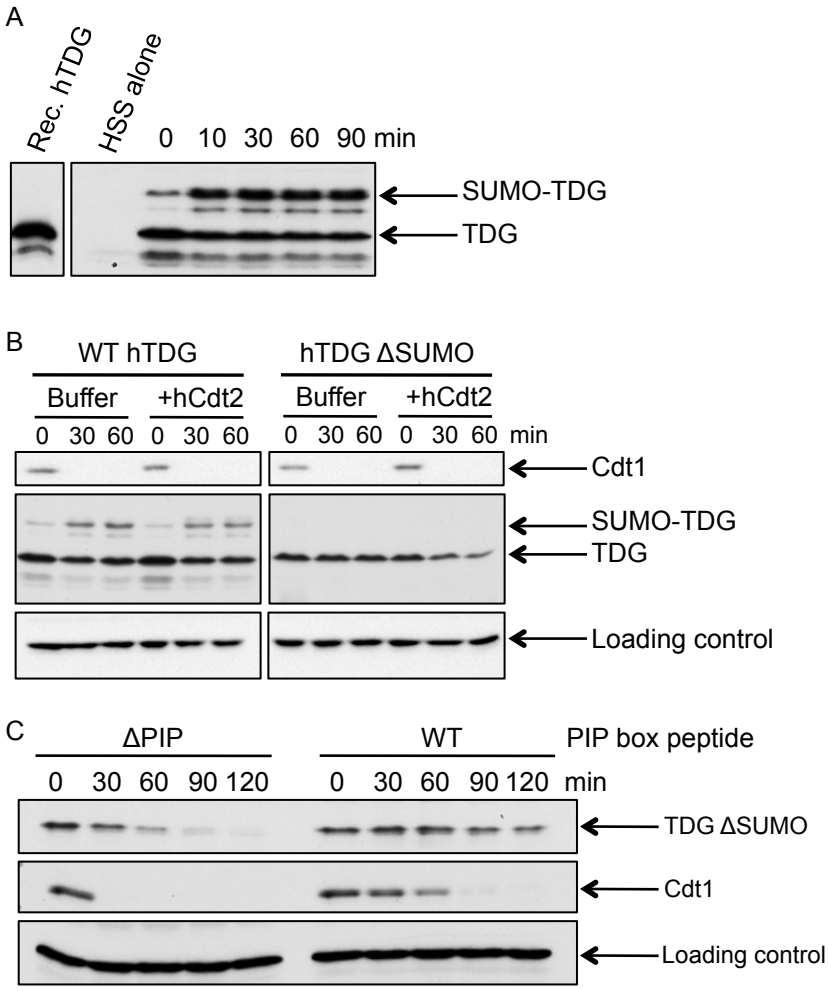



**Figure A3.3: Human TDG requires the addition of human Cdt2 to serve as a CRL4$^{Cdt2}$ substrate in *Xenopus* egg extract. A)** Human TDG is rapidly SUMOylated in HSS but is not destroyed in the presence of MMS-damaged plasmid. **B)** Human TDG ΔSUMO is destroyed in HSS upon the addition of an excess of recombinant human Cdt2. **C)** *Xenopus* TDG is destroyed slowly in HSS in a PCNA-dependent manner during DNA repair. The extract-specific, non-Cdt1 band detected by the Cdt1 antibody was used as a loading control.

# Appendix A4

# The generation of plasmids containing TDG substrates

## *A4.1 Introduction*

This appendix describes attempts to generate a new tool to study TDG function. Wildtype TDG is not destroyed during DNA replication in *Xenopus* egg extract. We therefore hypothesized that destruction might require TDG activity on the undamaged plasmid. During repair-dependent TDG destruction, abasic sites are likely present on the MMS-damaged plasmid, and TDG could bind directly to them. Thus, we tried several different approaches to generate a plasmid containing a site-specific TDG substrate. Ultimately, we used a plasmid containing a G-T mispair generated in the Takahashi laboratory at Osaka University in Japan (Chapter 2). Before using this plasmid construct, however, we tried several other approaches that are outlined here.

## *A4.2 Materials, Methods, and Results*

### A4.2.1 Incorporation of 5-FU via nick translation

TDG acts on 5-fluorouracil (5-FU) in DNA (Cortázar et al., 2007; Hardeland et al., 2000; 2003; Kunz et al., 2009). We therefore reasoned that if we incorporated 5-FU into a plasmid, TDG would bind to and catalytically act on that plasmid. We used a nick-translation approach to replace a small number of thymine moieties with 5-FU. In this approach, DNase I generates nicks within the plasmid, and then DNA Polymerase I performs strand displacement reactions to incorporate the dNTPs included in the reaction, generating short stretches of newly synthesized DNA on the plasmid. I included 5-FdUTP in place of dTTP in nick translation reactions on a small, undamaged plasmid (pCITE4a) and varied the timing of the reaction as well as the concentration of DNase I to control how many nicks occurred during nick translation. As a control, I included a small amount of α-$^{32}$P-dATP in a fraction of the total reaction to calculate the total incorporation nick translation for a given set of conditions. After the nick translation

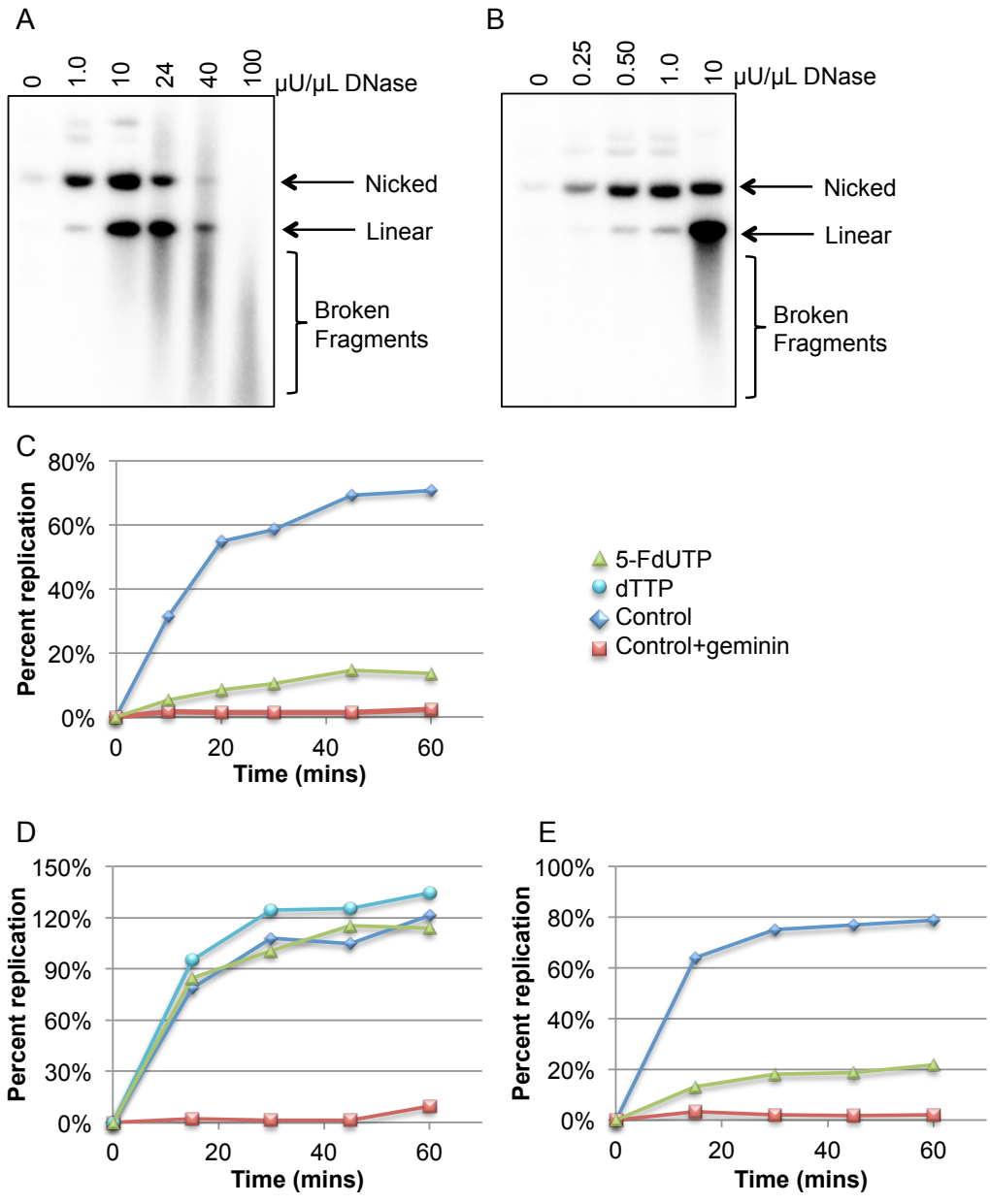
reaction was complete, reactions were stopped and treated with Proteinase K to remove DNase I and DNA Polymerase I from the DNA. Plasmids were recovered through phenol-chloroform extraction and ethanol precipitation.

Initial reactions set up based on nick translation kit protocols (Amersham, Invitrogen) led to high nucleotide incorporation. In these reactions, the concentration of DNase I varied from 0.25-100 μUnit DNase, and the total reaction time was 2 hours at 15°C. Reactions also contained 20 μM each dNTP, 5 mM $MgCl_2$, 10 mM β-mercaptoethanol, 50 mU/μL DNA Polymerase I, and 50mM Tris, pH 7.7. However, even at the lowest concentration of DNase tested, a fraction of the resultant plasmid was linear, indicating that DNase activity and/or strand displacement activity was extensive (Figure A4.1A, B). I titrated the concentration of DNase even further, but the plasmid did not replicate well in HSS/NPE (Figure A4.1C).

I modified the nick translation protocol to be more similar to one that was previously used in the Walter lab to label the parental strand during DNA replication (Knipscheer et al., 2009). In this modified nick translation protocol, the undamaged plasmid is first incubated with DNase (28.4 μU/μL) at room temperature for 1 minute before being diluted to a final concentration of 7.8 μU/μL with buffer containing dNTPs (20 μM each, final concentration) and DNA Polymerase I (250 mU/μL final concentration). Other conditions remained the same. I varied the time of the DNA Polymerase I reaction from 10-30 minutes at 15°C. After nick translation for 10 minutes, a negligible amount of incorporated dNTPs was detected (0.71%). Additionally, the ratio of nicked to supercoiled plasmid did not change upon the nick translation reaction, which was unexpected given the presence of a nicking enzyme (data not shown). This plasmid replicated well (Figure A4.1D), but we questioned whether it actually contained 5-FU. After a 30 minute nick translation reaction, however, we saw a greater amount of

**Figure A4.1: Nick translation incorporates 5-FU into an undamaged plasmid. A)** Nick translation was monitored through the addition of a trace amount of α-$^{32}$P-dATP. DNase concentrations were varied from 0-100 μUnits/μL during a two-hour nick translation reaction. **B)** DNase concentrations were varied from 0-10 μUnits/μL during a two-hour nick translation reaction. **C)** The plasmid that was generated through a two-hour nick-translation reaction led to weak DNA replication, even though the DNase concentration was very slow (2μUnits/μL). **D)** A 10-minute nick translation reaction with either 5-FdUTP or dTTP using the second nick-translation approach did not impair DNA replication. The 5-FU incorporation under these conditions was less than 1%. **E)** A 30-minute nick translation reaction using the second nick-translation approach impaired DNA replication. Total incorporation was 8.6%. A single figure legend is shown for (C)-(E).

**Figure A4.1** (continued)



dNTP incorporation (8.65%). Unfortunately, this plasmid did not replicate in HSS/NPE (Figure A4.1E). We terminated this approach at this point and moved on to the other approaches described below.

### A4.2.2 Incorporation of a short stretch of 5-FU-containing DNA via PCR and ligation

As an alternative approach to incorporate 5-FU into a plasmid, I included 5-FdUTP in a PCR reaction in place of dTTP and amplified a short region of pCITE4a. Notably, the high fidelity polymerase KOD Hi-Fi (Millipore) did not lead to any PCR products in reactions containing 5-FU, suggesting that it cannot incorporate the aberrant nucleotide. Taq polymerase generated the desired PCR product. TDG bound to this PCR product and also was active in a base release assay (Figure A4.2A, B), confirming that TDG acts on 5-FU, even when it is paired with adenine.

Attempts to use a restriction enzyme digest and ligation to insert this short PCR product containing 5-FU were unsuccessful. I chose to digest with the restriction enzyme NotI on both ends of the PCR product and in a single position on the plasmid because its recognition sequence contains only guanines and cytosines, so the presence of 5-FU would not impair its ability to cleave the insert. However, upon ligation, several higher migrating bands were visible on an agarose gel (Figure A4.2C), likely due to the insertion of multiple PCR products into a single plasmid. Though I could have changed the approach to use two distinct restriction enzymes, we decided to terminate this approach because we reasoned that the high number of aberrant bases would impair DNA replication, even in the absence of TDG.

### A4.2.3 Generation of plasmids containing site-specific G-U and G-T mispairs

The most successful approach for the generation of a plasmid containing a TDG

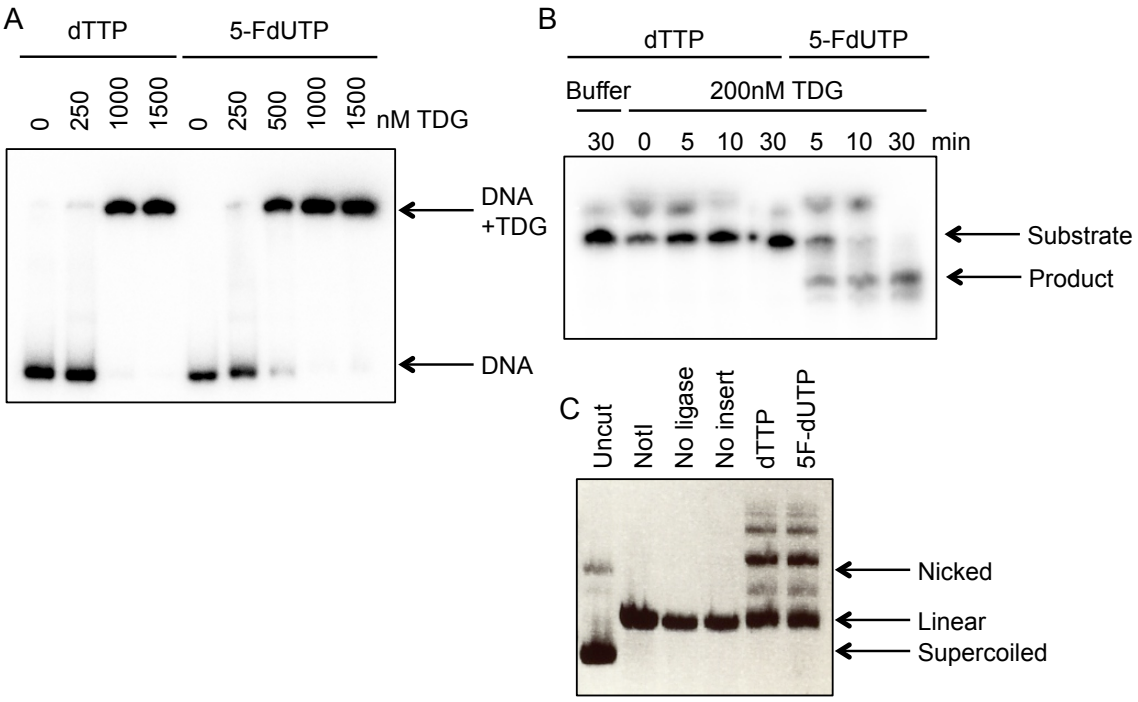



**Figure A4.2: A short PCR fragment containing 5-FU serves as a TDG substrate but was not successfully ligated into a plasmid. A)** TDG binds to a short PCR fragment containing standard nucleotides or 5-FU. Shown here is a native acrylamide gel. TDG binds the PCR product containing dTTP similarly to that containing 5F-dUTP. **B)** TDG activity assay on this same PCR product. TDG processed the 5-FU-containing PCR product to a smaller end labeled product. **C)** Ligation of the 5-FU-containing PCR product into pCITE4a results in numerous higher migrating bands, thought to be ligation products.

substrate involved generating a gapped plasmid and inserting a short oligo containing a site-specific mispair into the gap. This approach, based on a previously described method to generate a plasmid containing a site-specific abasic site (Matsumoto, 2006) successfully generated a plasmid containing a site-specific G-U mispair. This approach could be used in the future for the generation of mispaired plasmids or plasmids containing modified bases.

This approach requires an undamaged plasmid that contains a restriction enzyme site sandwiched between two nicking sites separated by 20-30 base pairs. The desired mispair should fall within the restriction enzyme recognition site. To start, the plasmid is digested with the nicking enzymes. The nicked plasmid is then heated in the presence of an excess of competitor oligo (the complement of the short fragment between the nicking sites) to generate a gapped plasmid. The resulting mixture is then digested with the restriction enzyme that cleaves between the nicking sites. The digest should linearize the nicked plasmid but not the gapped plasmid, making them separable on an agarose gel. After purification of the gapped plasmid through gel electrophoresis and electroelution, a short oligo that complements the now single-stranded portion of the plasmid except at the desired site-specific mispair location is annealed to the gapped plasmid. Finally, the plasmid is ligated, and the closed circular plasmid is gel purified from an agarose gel containing 3.6 μM chloroquine. This final purification step could be modified to involve purification through a cesium chloride gradient to scale up the reaction. Many of the details regarding concentration of the nicking enzyme and timing of the nicking digest were determined by Thomas Graham of the Walter and Loparo labs.

I attempted this approach on three separate parent plasmids. I first tried to insert the mispair into pCtrl, the plasmid that many Walter lab members use as a control to study interstrand crosslink repair (Knipscheer et al., 2009; Long et al., 2011; Räschle et al.,

2008). I inserted two Nb.BbvcI nicking sites on either side of a SapI recognition site. Notably, pCtrl was not compatible with the QuikChange mutagenesis kit; I performed mutagenesis on a small region of pCtrl in the context of a pDonr backbone. Dr. David Long, a post-doctoral fellow in the Walter lab, generated this construct. This SapI recognition site also contains an interstrand crosslink in pICL (Knipscheer et al., 2009; Long et al., 2011; Räschle et al., 2008), which was ideal since reagents were already available in the lab to follow pICL replication and repair. However, after multiple failed attempts to ligate this plasmid, I switched to using pBluescript.

I used two variations of pBluescript: pBSRon, generated by Dr. Ron Lebofsky and Dr. Courtney Havens, previously of the Walter lab, to contain two Nb.BtsI nicking sites sandwiching an NcoI restriction enzyme site, and pBSTom, generated by Thomas Graham, of the Walter and Loparo labs, to contain two Nb.BbvCI nicking sites sandwiching an EcoRI restriction enzyme site. I generated closed circular products from both of these parental plasmids, but I found that NcoI displays some activity on the plasmid containing a G-U mispair, so I ultimately used pBSTom exclusively.

TDG acted on this mispair in the context of the plasmid (Figure A4.3A), though activity required a higher concentration than TDG activity on a short oligo of the same sequence, likely because TDG associates with non-specific DNA on the plasmid. However, the G-U mispair was repaired rapidly in extract to regenerate the EcoRI recognition site (Figure A4.3B), likely due to uracil DNA glycosylase (UDG) activity in extract. pBSTom (G-U) replicated normally (Figure A4.3C). We next wanted to generate a plasmid containing a G-T mispair. However, the designed G-U mispair was not in the context of a CpG sequence, and TDG prefers this sequence context. Recombinant *Xenopus* TDG did not act on a G-T mispair in the designed sequence. There is an EcoRV recognition site adjacent to the EcoRI recognition site that would allow the generation of a G-T mispair

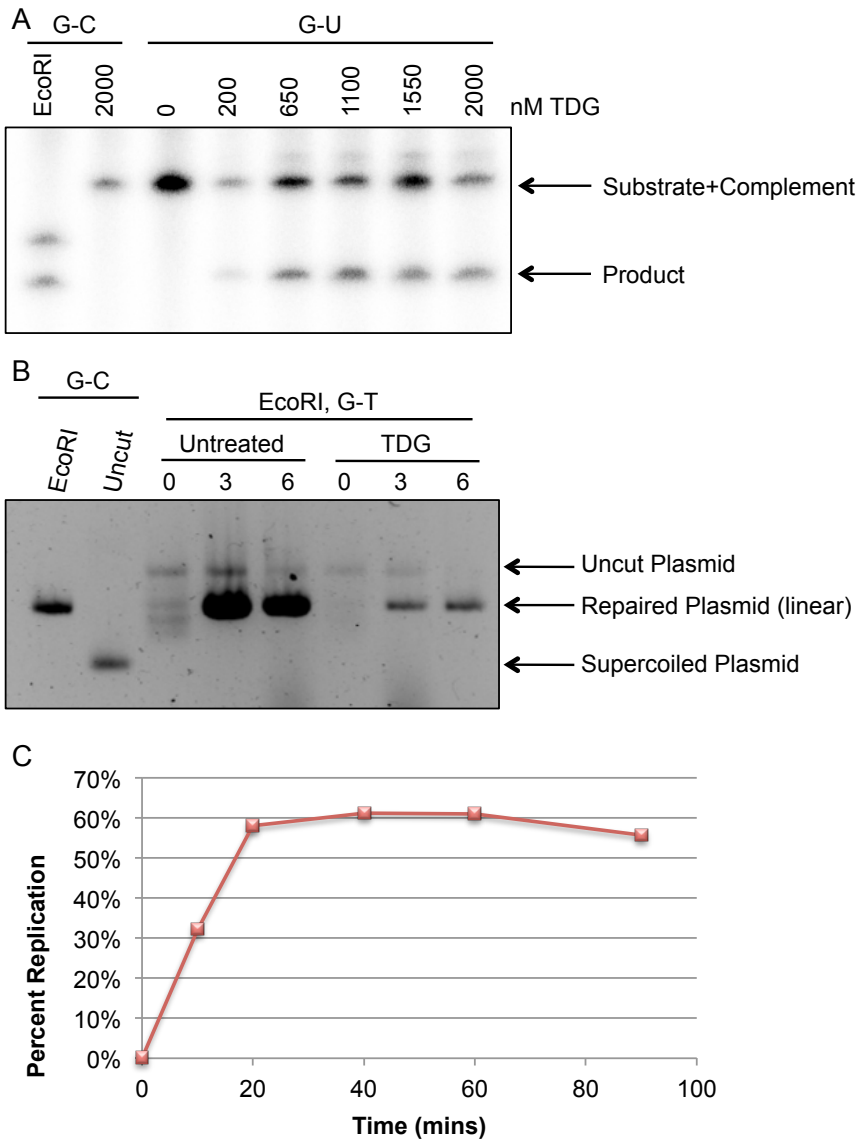



**Figure A4.3: The gapped plasmid method of generating a G-U mispaired plasmid was successful. A)** TDG is active on the G-U mispair in the context of the plasmid. This reaction was performed as described in Chapter 2, Figure 2.1, except the mispair was adjacent to an EcoRI site instead of a NotI site. An EcoRI digest is shown to indicate the expected size of the final product. **B)** The G-U mispair is rapidly repaired in extract, regardless of pre-binding with TDG. Repair was measured by susceptibility to an EcoRI digest. **C)** The G-U mispaired plasmid replicates in extract. Quantification of DNA synthesis is shown.

within a CpG sequence. However, we opted to outsource the production of a G-T mispair-containing plasmid to the Takahashi lab at this point.

Although we were not able to use this approach to generate a DNA template for use in extract experiments, this approach has now been optimized and could easily be modified for future use in the Walter lab. Gel electroelution of the gapped plasmid led to the recovery of 40-60% of the original input (approximately 85-95% of the gapped fraction). Once the gapped plasmid is generated, insertion of the mispair and subsequent purification is simple and has been previously outlined (Matsumoto, 2006). I chose to use electroelution as the final means of plasmid purification, but cesium chloride gradient purification might be even more successful for larger-scale production.

# *List of Abbreviations*

| | |
|---|---|
| 5-FU | 5-fluorouracil |
| 5caC | 5-carboxylcytosine |
| 5fC | 5-formylcytosine |
| 5hmC | 5-hydroxymethylcytosine |
| 5hmU | 5-hydroxymethyluracil |
| 5mC | 5-methylcytosine |
| AID | Activation-induced cytosine deaminase |
| AP | Apurinic |
| ATM | Ataxia-telangiectasia mutated |
| ATR | ATM and Rad3-related |
| BER | Base excision repair |
| Cdk | Cyclin-dependent kinase |
| Chk1 | Checkpoint kinase 1 |
| CI | Catalytically inactive |
| CMG | Cdc45/MCM2-7/GINS helicase |
| CRL | Cullin-RING-based ubiquitin ligase |
| CSF | Cytostatic factor |
| DBD | DNA binding domain |
| Dnmt | DNA methyltransferase |
| ELB | Egg lysis buffer |
| HCG | Human chorionic gonadotropin |
| IDCL | Interdomain connector loop |
| MMR | Marc's modified ringer solution |
| M phase | Mitosis |
| NER | Nucleotide excision repair |
| NLS | Nuclear localization sequence |
| ORC | Origin recognition complex |
| PCNA | Proliferating cell nuclear antigen |
| $PCNA^{SC}$ | Single-chain PCNA |
| PIP | PCNA-interacting peptide |
| PMSG | Pregnant mare serum gonadotropin |
| Pol | Polymerase |
| pre-IC | Pre-initiation complex |
| pre-RC | Pre-replication complex |
| RFC | Replication factor C |
| RPA | Replication protein A |
| S phase | Synthesis Phase |
| ssDNA | Single-stranded DNA |
| TDG | Thymine DNA glycosylase |
| Tet | Ten eleven translocation |
| TLS | Translesion synthesis |
| UNG | Uracil-N-glycosylase |
| XPV | Xeroderma pigmentosum variant |

# *References*


Abbas, T., and Dutta, A. (2011). $CRL4^{Cdt2}$: master coordinator of cell cycle progression and genome stability. Cell Cycle *10*, 241–249.

Abbas, T., Shibata, E., Park, J., Jha, S., Karnani, N., and Dutta, A. (2010). $CRL4^{Cdt2}$ regulates cell proliferation and histone gene expression by targeting PR-Set7/Set8 for degradation. Molecular Cell *40*, 9–21.

Acharya, N., Yoon, J.-H., Gali, H., Unk, I., Haracska, L., Johnson, R.E., Hurwitz, J., Prakash, L., and Prakash, S. (2008). Roles of PCNA-binding and ubiquitin-binding domains in human DNA polymerase eta in translesion DNA synthesis. Proc. Natl. Acad. Sci. *105*, 17724–17729.

Albertella, M.R. (2005). A role for polymerase η in the cellular tolerance to cisplatin-induced damage. Cancer Research *65*, 9799–9806.

Andreassen, P.R. (2005). DNA damage responses and their many interactions with the replication fork. Carcinogenesis *27*, 883–892.

Arias, E.E., and Walter, J.C. (2005). Replication-dependent destruction of Cdt1 limits DNA replication to a single round per cell cycle in *Xenopus* egg extracts. Genes & Development *19*, 114–126.

Arias, E.E., and Walter, J.C. (2007). Strength in numbers: preventing rereplication via multiple mechanisms in eukaryotic cells. Genes & Development *21*, 497–518.

Arias, E.E., and Walter, J.C. (2004). Initiation of DNA replication in *Xenopus* egg extracts. Front. Biosci. *9*, 3029–3045.

Arias, E.E., and Walter, J.C. (2006). PCNA functions as a molecular platform to trigger Cdt1 destruction and prevent re-replication. Nat Cell Biol *8*, 84–90.

Baba, D., Maita, N., Jee, J.-G., Uchimura, Y., Saitoh, H., Sugasawa, K., Hanaoka, F., Tochio, H., Hiroaki, H., and Shirakawa, M. (2005). Crystal structure of thymine DNA glycosylase conjugated to SUMO-1. Nature *435*, 979–982.

Barreto, G., Schäfer, A., Marhold, J., Stach, D., Swaminathan, S.K., Handa, V., Döderlein, G., Maltry, N., Wu, W., Lyko, F., et al. (2007). Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. Nature *445*, 671–675.

Beck, D.B., Oda, H., Shen, S.S., and Reinberg, D. (2012). PR-Set7 and H4K20me1: at the crossroads of genome integrity, cell cycle, chromosome condensation, and transcription. Genes & Development *26*, 325–337.

Blitz, I.L., Andelfinger, G., and Horb, M.E. (2006). Germ layers to organs: Using *Xenopus* to study "later" development. Seminars in Cell & Developmental Biology *17*, 133–145.

Blow, J.J. (1993). Preventing re-replication of DNA in a single cell cycle: evidence for a

replication licensing factor. The Journal of Cell Biology *122*, 993–1002.

Blow, J.J., and Laskey, R.A. (1988). A role for the nuclear envelope in controlling DNA replication within the cell cycle. Nature *332*, 546–548.

Blow, J.J., Ge, X.Q., and Jackson, D.A. (2011). How dormant origins promote complete genome replication. Trends in Biochemical Sciences *36*, 405–414.

Bouvet, P., Dimitrov, S., and Wolffe, A.P. (1994). Specific regulation of *Xenopus* chromosomal 5S rRNA gene transcription in vivo by histone H1. Genes & Development *8*, 1147–1159.

Bowes, J.B., Snyder, K.A., Segerdell, E., Gibb, R., Jarabek, C., Noumen, E., Pollet, N., and Vize, P.D. (2007). Xenbase: a *Xenopus* biology and genomics resource. Nucleic Acids Research *36*, D761–D767.

Bullock, S.K., Kaufmann, W.K., and Cordeiro-Stone, M. (2001). Enhanced S phase delay and inhibition of replication of an undamaged shuttle vector in UVC-irradiated xeroderma pigmentosum variant. Carcinogenesis *22*, 233–241.

Burgers, P.M.J. (2008). Polymerase dynamics at the eukaryotic DNA replication fork. Journal of Biological Chemistry *284*, 4041–4045.

Byun, T.S. (2005). Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint. Genes & Development *19*, 1040–1052.

Carey, M.F., Peterson, C.L., and Smale, S.T. (2009). Dignam and Roeder nuclear extract preparation. Cold Spring Harbor Protocols *2009*, pdb.prot5330–pdb.prot5330.

Centore, R.C., Havens, C.G., Manning, A.L., Li, J.-M., Flynn, R.L., Tse, A., Jin, J., Dyson, N.J., Walter, J.C., and Zou, L. (2010). $CRL4^{Cdt2}$-mediated destruction of the histone methyltransferase Set8 prevents premature chromatin compaction in S phase. Molecular Cell *40*, 22–33.

Chen, Z.X., and Riggs, A.D. (2011). DNA methylation and demethylation in mammals. Journal of Biological Chemistry *286*, 18347–18353.

Cimmino, L., Abdel-Wahab, O., Levine, R.L., and Aifantis, I. (2011). TET family proteins and their role in stem cell differentiation and transformation. Stem Cell *9*, 193–204.

Clute, P., and Masui, Y. (1997). Microtubule dependence of chromosome cycles in *Xenopus* laevis blastomeres under the influence of a DNA synthesis inhibitor, aphidicolin. Dev. Biol. *185*, 1–13.

Cocker, J.H., Piatti, S., Santocanale, C., Nasmyth, K., and Diffley, J.F. (1996). An essential role for the Cdc6 protein in forming the pre-replicative complexes of budding yeast. Nature *379*, 180–182.

Conn, C.W., Lewellyn, A.L., and Maller, J.L. (2004). The DNA damage checkpoint in embryonic cell cycles is dependent on the DNA-to-cytoplasmic ratio. Developmental Cell

*7*, 275–281.

Cortázar, D., Kunz, C., Saito, Y., Steinacher, R., and Schär, P. (2007). The enigmatic thymine DNA glycosylase. DNA Repair *6*, 489–504.

Cortázar, D., Kunz, C., Selfridge, J., Lettieri, T., Saito, Y., MacDougall, E., Wirz, A., Schuermann, D., Jacobs, A.L., Siegrist, F., et al. (2011). Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. Nature *470*, 419–423.

Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., Cigliano, A., Le Coz, M., Devarajan, K., Wessels, A., Soprano, D., et al. (2011). Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. Cell *146*, 67–79.

Cvetic, C., and Walter, J.C. (2005). Eukaryotic origins of DNA replication: could you please be more specific? Seminars in Cell & Developmental Biology *16*, 343–353.

Dalton, S.R., and Bellacosa, A. (2012). DNA demethylation by TDG. Epigenomics *4*, 459–467.

Dasso, M., and Newport, J.W. (1990). Completion of DNA replication is monitored by a feedback system that controls the initiation of mitosis in vitro: studies in *Xenopus*. Cell *61*, 811–823.

Davidson, L.A., Hoffstrom, B.G., Keller, R., and DeSimone, D.W. (2002). Mesendoderm extension and mantle closure in *Xenopus laevis* gastrulation: Combined roles for integrin α5β1, fibronectin, and tissue geometry. Dev. Biol. *242*, 109–129.

De Biasio, A., Sánchez, R., Prieto, J., Villate, M., Campos-Olivas, R., and Blanco, F.J. (2011). Reduced stability and increased dynamics in the human proliferating cell nuclear antigen (PCNA) relative to the yeast homolog. PLoS ONE *6*, e16600.

Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Research *11*, 1475–1489.

Durando, M., Tateishi, S., and Vaziri, C. (2013). A non-catalytic role of DNA polymerase η in recruiting Rad18 and promoting PCNA monoubiquitination at stalled replication forks. Nucleic Acids Research *41*, 3079–3093.

Eissenberg, J.C., Ayyagari, R., Gomes, X.V., and Burgers, P.M. (1997). Mutations in yeast proliferating cell nuclear antigen define distinct sites for interaction with DNA polymerase delta and DNA polymerase epsilon. Molecular and Cellular Biology *17*, 6367–6378.

Faber, J., and Nieuwkoop, P.D. (1994). Normal table of *Xenopus laevis* (Daudin).

Freudenthal, B.D., Brogie, J.E., Gakhar, L., Kondratick, C.M., and Washington, M.T. (2011). Crystal Structure of SUMO-Modified Proliferating Cell Nuclear Antigen. J. Mol. Biol. *406*, 9–17.

Fritz, E.L., and Papavasiliou, F.N. (2010). Cytidine deaminases: AIDing DNA demethylation? Genes & Development *24*, 2107–2114.

Gillespie, P.J., Gambus, A., and Blow, J.J. (2012). Preparation and use of *Xenopus* egg extracts to study DNA replication and chromatin associated proteins. Methods *57*, 203–213.

Gomes, X.V., and Burgers, P.M. (2000). Two modes of FEN1 binding to PCNA regulated by DNA. The EMBO Journal *19*, 3811–3821.

Grayson, D.R., and Guidotti, A. (2012). The dynamics of DNA methylation in schizophrenia and related psychiatric disorders. Neuropsychopharmacology *38*, 138–166.

Greenwood, J., and Gautier, J. (2005). From oogenesis through gastrulation: developmental regulation of apoptosis. Seminars in Cell & Developmental Biology *16*, 215–224.

Gregory, D.J., Mikhaylova, L., and Fedulov, A.V. (2012). Selective DNA demethylation by fusion of TDG with a sequence-specific DNA-binding domain. Epigenetics *7*, 344–349.

Gulbis, J.M., Kelman, Z., Hurwitz, J., O'Donnell, M., and Kuriyan, J. (1996). Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA. Cell *87*, 297–306.

Guo, J.U., Su, Y., Zhong, C., Ming, G.-L., and Song, H. (2011). Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. Cell *145*, 423–434.

Hardeland, U., Bentele, M., Jiricny, J., and Schär, P. (2000). Separating substrate recognition from base hydrolysis in human thymine DNA glycosylase by mutational analysis. Journal of Biological Chemistry *275*, 33449–33456.

Hardeland, U., Bentele, M., Jiricny, J., and Schär, P. (2003). The versatile thymine DNA-glycosylase: a comparative characterization of the human, *Drosophila* and fission yeast orthologs. Nucleic Acids Research *31*, 2261–2271.

Hardeland, U., Kunz, C., Focke, F., Szadkowski, M., and Schär, P. (2007). Cell cycle regulation as a mechanism for functional separation of the apparently redundant uracil DNA glycosylases TDG and UNG2. Nucleic Acids Research *35*, 3859–3867.

Hardeland, U., Steinacher, R., Jiricny, J., and Schär, P. (2002). Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover. The EMBO Journal *21*, 1456–1464.

Hashimoto, H., Liu, Y., Upadhyay, A.K., Chang, Y., Howerton, S.B., Vertino, P.M., Zhang, X., and Cheng, X. (2012). Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. Nucleic Acids Research *40*, 4841–4849.

Havens, C.G., and Walter, J.C. (2009). Docking of a specialized PIP box onto chromatin-bound PCNA creates a degron for the ubiquitin ligase $CRL4^{Cdt2}$. Molecular Cell *35*, 93–

104.

Havens, C.G., and Walter, J.C. (2011). Mechanism of CRL4$^{Cdt2}$, a PCNA-dependent E3 ubiquitin ligase. Genes & Development *25*, 1568–1582.

Havens, C.G., Shobnam, N., Guarino, E., Centore, R.C., Zou, L., Kearsey, S.E., and Walter, J.C. (2012). Direct role for proliferating cell nuclear antigen in substrate recognition by the E3 ubiquitin ligase CRL4$^{Cdt2}$. Journal of Biological Chemistry *287*, 11410–11421.

He, X.-J., Chen, T., and Zhu, J.-K. (2011a). Regulation and function of DNA methylation in plants and animals. Nature Publishing Group *21*, 442–465.

He, Y.F., Li, B.Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, L., et al. (2011b). Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science *333*, 1303–1307.

Heller, R.C., Kang, S., Lam, W.M., Chen, S., Chan, C.S., and Bell, S.P. (2011). Eukaryotic origin-dependent DNA replication *in vitro* reveals sequential action of DDK and S-CDK kinases. Cell *146*, 80–91.

Hensey, C., and Gautier, J. (1998). Programmed cell death during *Xenopus* development: a spatio-temporal analysis. Dev. Biol. *203*, 36–48.

Holway, A.H., Kim, S.-H., La Volpe, A., and Michael, W.M. (2006). Checkpoint silencing during the DNA damage response in *Caenorhabditis elegans* embryos. The Journal of Cell Biology *172*, 999–1008.

Hu, J., McCall, C.M., Ohta, T., and Xiong, Y. (2004). Targeted ubiquitination of CDT1 by the DDB1–CUL4A–ROC1 ligase in response to DNA damage. Nat Cell Biol *6*, 1003–1009.

Hua, X.H., Yan, H., and Newport, J. (1997). A role for Cdk2 kinase in negatively regulating DNA replication during S phase of the cell cycle. The Journal of Cell Biology *137*, 183–192.

Ilves, I., Petojevic, T., Pesavento, J.J., and Botchan, M.R. (2010). Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. Molecular Cell *37*, 247–258.

Indiani, C., McInerney, P., Georgescu, R., Goodman, M.F., and O'Donnell, M. (2005). A sliding-clamp toolbelt binds high- and low-fidelity DNA polymerases simultaneously. Molecular Cell *19*, 805–815.

Inoue, A., and Zhang, Y. (2011). Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. Science *334*, 194–194.

Inoue, A., Shen, L., Dai, Q., He, C., and Zhang, Y. (2011). Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development. Nature Publishing Group *21*, 1670–1676.

Iqbal, K., Jin, S.-G., Pfeifer, G.P., and Szabó, P.E. (2011). Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. Proc. Natl. Acad. Sci. *108*, 3642–3647.

Ise, T., Nagatani, G., Imamura, T., Kato, K., Takano, H., Nomoto, M., Izumi, H., Ohmori, H., Okamoto, T., Ohga, T., et al. (1999). Transcription factor Y-box binding protein 1 binds preferentially to cisplatin-modified DNA and interacts with proliferating cell nuclear antigen. Cancer Research *59*, 342–346.

Ito, S., Shen, L., Dai, Q., Wu, S.C., Collins, L.B., Swenberg, J.A., He, C., and Zhang, Y. (2011). Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science *333*, 1300–1303.

Jeltsch, A. (2006). On the enzymatic properties of Dnmt1: specificity, processivity, mechanism of linear diffusion and allosteric regulation of the enzyme. Epigenetics *1*, 63–66.

Ji, H., Ehrlich, L.I.R., Seita, J., Murakami, P., Doi, A., Lindau, P., Lee, H., Aryee, M.J., Irizarry, R.A., Kim, K., et al. (2010). Comprehensive methylome map of lineage commitment from haematopoietic progenitors. Nature *467*, 338–342.

Jin, J., Arias, E.E., Chen, J., Harper, J.W., and Walter, J.C. (2006). A family of diverse Cul4-Ddb1-interacting proteins includes Cdt2, which is required for S phase destruction of the replication factor Cdt1. Molecular Cell *23*, 709–721.

Johnson, R.E., Haracska, L., Prakash, S., and Prakash, L. (2001). Role of DNA polymerase η in the bypass of a (6-4) TT photoproduct. Molecular and Cellular Biology *21*, 3558–3563.

Johnson, R.E., Prakash, S., and Prakash, L. (1999). Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase, Polη. Science *283*, 1001–1004.

Joseph, R.E., and Andreotti, A.H. (2008). Bacterial expression and purification of Interleukin-2 Tyrosine kinase: Single step separation of the chaperonin impurity. Protein Expression and Purification *60*, 194–197.

Kamileri, I., Karakasilioti, I., and Garinis, G.A. (2012). Nucleotide excision repair: new tricks with old bricks. Trends in Genetics *28*, 566–573.

Kappas, N.C., Savage, P., Chen, K.C., Walls, A.T., and Sible, J.C. (2000). Dissection of the XChk1 signaling pathway in *Xenopus* laevis embryos. Mol. Biol. Cell *11*, 3101–3108.

Kim, B.J., and Lee, H. (2012). Lys-110 is essential for targeting PCNA to replication and repair foci, and the K110A mutant activates apoptosis. Biology of the Cell *100*, 675–686.

Kim, D.H., Budhavarapu, V.N., Herrera, C.R., Nam, H.W., Kim, Y.S., and Yew, P.R. (2010). The $CRL4^{Cdt2}$ ubiquitin ligase mediates the proteolysis of cyclin-dependent kinase inhibitor Xic1 through a direct association with PCNA. Molecular and Cellular Biology *30*, 4120–4133.

Kim, J., and Kipreos, E.T. (2008). Control of the Cdc6 replication licensing factor in

metazoa: the role of nuclear export and the CUL4 ubiquitin ligase. Cell Cycle *7*, 146–150.

Kim, S.-H., and Michael, W.M. (2008). Regulated proteolysis of DNA polymerase η during the DNA-damage response in *C. elegans*. Molecular Cell *32*, 757–766.

Kim, Y., Starostina, N.G., and Kipreos, E.T. (2008). The $CRL4^{Cdt2}$ ubiquitin ligase targets the degradation of $p21^{Cip1}$ to control replication licensing. Genes & Development *22*, 2507–2519.

Kim, Y.-J., and Wilson, D.M. (2012). Overview of base excision repair biochemistry. Curr Mol Pharmacol *5*, 3–13.

Kimelman, D., Kirschner, M., and Scherson, T. (1987). The events of the midblastula transition in *Xenopus* are regulated by changes in the cell cycle. Cell *48*, 399–407.

Knipscheer, P., Raschle, M., Smogorzewska, A., Enoiu, M., Ho, T.V., Scharer, O.D., Elledge, S.J., and Walter, J.C. (2009). The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. Science *326*, 1698–1701.

Kochaniak, A.B., Habuchi, S., Loparo, J.J., Chang, D.J., Cimprich, K.A., Walter, J.C., and van Oijen, A.M. (2009). Proliferating cell nuclear antigen uses two distinct modes to move along DNA. J. Biol. Chem. *284*, 17700–17710.

Kress, C., Thomassin, H., and Grange, T. (2006). Active cytosine demethylation triggered by a nuclear receptor involves DNA strand breaks. Proc. Natl. Acad. Sci. *103*, 11112–11117.

Krokan, H.E., Drabløs, F., and Slupphaug, G. (2002). Uracil in DNA—occurrence, consequences and repair. Oncogene *21*, 8935–8948.

Kumagai, A., and Dunphy, W.G. (1999). Binding of 14-3-3 proteins and nuclear export control the intracellular localization of the mitotic inducer Cdc25. Genes & Development *13*, 1067–1072.

Kumagai, A., Guo, Z., Emami, K.H., Wang, S.X., and Dunphy, W.G. (1998). The *Xenopus* Chk1 protein kinase mediates a caffeine-sensitive pathway of checkpoint control in cell-free extracts. The Journal of Cell Biology *142*, 1559–1569.

Kunz, C., Focke, F., Saito, Y., Schuermann, D., Lettieri, T., Selfridge, J., and Schär, P. (2009). Base excision by thymine DNA glycosylase mediates DNA-directed cytotoxicity of 5-fluorouracil. Plos Biol *7*, e91.

Leach, P.T., Poplawski, S.G., Kenney, J.W., Hoffman, B., Liebermann, D.A., Abel, T., and Gould, T.J. (2012). Gadd45b knockout mice exhibit selective deficits in hippocampus-dependent long-term memory. Learning & Memory *19*, 319–324.

Lebofsky, R., Takahashi, T., and Walter, J.C. (2009). DNA replication in nucleus-free *Xenopus* egg extracts. Methods Mol. Biol. *521*, 229–252.

Li, A., and Blow, J.J. (2004). Non-proteolytic inactivation of geminin requires CDK-dependent ubiquitination. Nat Cell Biol *6*, 8–267.

Li, Y.Q., Zhou, P.Z., Zheng, X.D., Walsh, C.P., and Xu, G.L. (2006). Association of Dnmt3a and thymine DNA glycosylase links DNA methylation with base-excision repair. Nucleic Acids Research *35*, 390–400.

Liu, C., Powell, K.A., Mundt, K., Wu, L., Carr, A.M., and Caspari, T. (2003). Cop9/signalosome subunits and Pcu4 regulate ribonucleotide reductase by both checkpoint-dependent and -independent mechanisms. Genes & Development *17*, 1130–1140.

Liu, J., Grimison, B., and Maller, J.L. (2007). New insight into metaphase arrest by cytostatic factor: from establishment to release. Oncogene *26*, 1286–1289.

Long, D.T., Raschle, M., Joukov, V., and Walter, J.C. (2011). Mechanism of RAD51-dependent DNA interstrand cross-link repair. Science *333*, 84–87.

Longley, D.B., Harkin, D.P., and Johnston, P.G. (2003). 5-Fluorouracil: mechanisms of action and clinical strategies. Nat Rev Cancer *3*, 330–338.

Lopes, M., Foiani, M., and Sogo, J.M. (2006). Multiple mechanisms control chromosome integrity after replication fork uncoupling and restart at irreparable UV lesions. Molecular Cell *21*, 15–27.

Lupardus, P.J., Byun, T., Yee, M.-C., Hekmat-Nejad, M., and Cimprich, K.A. (2002). A requirement for replication in activation of the ATR-dependent DNA damage checkpoint. Genes & Development *16*, 2327–2332.

MacDougall, C.A., Byun, T.S., Van, C., Yee, M.C., and Cimprich, K.A. (2007). The structural determinants of checkpoint activation. Genes & Development *21*, 898–903.

Maiorano, D., Moreau, J., and Méchali, M. (2000). XCDT1 is required for the assembly of pre-replicative complexes in *Xenopus laevis*. Nature *404*, 622–625.

Maiti, A., and Drohat, A.C. (2011). Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. Journal of Biological Chemistry *286*, 35334–35338.

Maiti, A., Morgan, M.T., Pozharski, E., and Drohat, A.C. (2008). Crystal structure of human thymine DNA glycosylase bound to DNA elucidates sequence-specific mismatch recognition. Proc. Natl. Acad. Sci. *105*, 8890–8895.

Makarova, K.S., Koonin, E.V., and Kelman, Z. (2012). The CMG (CDC45/RecJ, MCM, GINS) complex is a conserved component of the DNA replication system in all archaea and eukaryotes. Biology Direct *7*, 7.

Masai, H., Matsumoto, S., You, Z., Yoshizawa-Sugata, N., and Oda, M. (2010). Eukaryotic chromosome DNA replication: where, when, and how? Annu. Rev. Biochem. *79*, 89–130.

Masui, Y., and Markert, C.L. (1971). Cytoplasmic control of nuclear behavior during meiotic maturation of frog oocytes. J. Exp. Zool. *177*, 129–145.

Masutani, C., Kusumoto, R., Yamada, A., Dohmae, N., Yokoi, M., Yuasa, M., Araki, M., Iwai, S., Takio, K., and Hanaoka, F. (1999). The XPV (xeroderma pigmentosum variant) gene encodes human DNA polymerase η. Nature *399*, 700–704.

Matsumoto, Y. (2006). Base excision repair in mammalian cells. Methods Mol. Biol. *314*, 365–375.

Maunakea, A.K., Chepelev, I., Zhao, K., and Bruneau, B. (2010). Epigenome mapping in normal and disease states. Circulation Research *107*, 327–339.

Mayer, W., Niveleau, A., Walter, J.C., Fundele, R., and Haaf, T. (2000). Demethylation of the zygotic paternal genome. Nature *403*, 501–502.

McCulloch, S.D., and Kunkel, T.A. (2008). The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. Cell Res *18*, 148–161.

McGarry, T.J., and Kirschner, M.W. (1998). Geminin, an inhibitor of DNA replication, is degraded during mitosis. Cell *93*, 1043–1053.

McNally, R., Bowman, G.D., Goedken, E.R., O'Donnell, M., and Kuriyan, J. (2010). Analysis of the role of PCNA-DNA contacts during clamp loading. BMC Struct Biol *10*, 3.

Mohan, R.D., Litchfield, D.W., Torchia, J., and Tini, M. (2010). Opposing regulatory roles of phosphorylation and acetylation in DNA mispair processing by thymine DNA glycosylase. Nucleic Acids Research *38*, 1135–1148.

Moldovan, G.-L., Pfander, B., and Jentsch, S. (2007). PCNA, the maestro of the replication fork. Cell *129*, 665–679.

Morgan, H.D., Dean, W., Coker, H.A., Reik, W., and Petersen-Mahrt, S.K. (2004). Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. Journal of Biological Chemistry *279*, 52353–52360.

Morgan, M.T., Bennett, M.T., and Drohat, A.C. (2007). Excision of 5-halogenated uracils by human thymine DNA glycosylase. Robust activity for DNA contexts other than CpG. J. Biol. Chem. *282*, 27578–27586.

Mortusewicz, O., Roth, W., Li, N., Cardoso, M.C., Meisterernst, M., and Leonhardt, H. (2008). Recruitment of RNA polymerase II cofactor PC4 to DNA damage sites. The Journal of Cell Biology *183*, 769–776.

Murray, A. (1995). Cyclin ubiquitination: the destructive end of mitosis. Cell *81*, 149–152.

Murray, A.W., and Kirschner, M.W. (1989). Cyclin synthesis drives the early embryonic cell cycle. Nature *339*, 275–280.

Nabel, C.S., and Kohli, R.M. (2011). Demystifying DNA demethylation. Science *333*, 1229–1230.

Naryzhny, S.N., DeSouza, L.V., Siu, K.W.M., and Lee, H. (2006). Characterization of the

human proliferating cell nuclear antigen physico-chemical properties: aspects of double trimer stability. Biochem. Cell Biol. *84*, 669–676.

Naryzhny, S.N., Zhao, H., and Lee, H. (2005). Proliferating cell nuclear antigen (PCNA) may function as a double homotrimer complex in the mammalian cell. Journal of Biological Chemistry *280*, 13888–13894.

Newport, J. (1987). Nuclear reconstitution in vitro: stages of assembly around protein-free DNA. Cell *48*, 205–217.

Newport, J., and Kirschner, M. (1982a). A major developmental transition in early *Xenopus* embryos: I. characterization and timing of cellular changes at the midblastula stage. Cell *30*, 675–686.

Newport, J., and Kirschner, M. (1982b). A major developmental transition in early *Xenopus* embryos: II. Control of the onset of transcription. Cell *30*, 687–696.

Nick McElhinny, S.A., Gordenin, D.A., Stith, C.M., Burgers, P.M.J., and Kunkel, T.A. (2008). Division of labor at the eukaryotic replication fork. Molecular Cell *30*, 137–144.

Niehrs, C. (2009). Active DNA demethylation and DNA repair. Differentiation *77*, 1–11.

Niehrs, C., and Schäfer, A. (2012). Active DNA demethylation by Gadd45 and DNA repair. Trends Cell Biol. *22*, 220–227.

Okada, Y., Yamagata, K., Hong, K., Wakayama, T., and Zhang, Y. (2010). A role for the elongator complex in zygotic paternal genome demethylation. Nature *463*, 554–558.

Oku, T., Ikeda, S., Sasaki, H., Fukuda, K., Morioka, H., Ohtsuka, E., Yoshikawa, H., and Tsurimoto, T. (1998). Functional sites of human PCNA which interact with p21 (Cip1/Waf1), DNA polymerase delta and replication factor C. Genes Cells *3*, 357–369.

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., Fundele, R., Dean, W., Reik, W., and Walter, J. (2000). Active demethylation of the paternal genome in the mouse zygote. Curr. Biol. *10*, 475–478.

Popp, C., Dean, W., Feng, S., Cokus, S.J., Andrews, S., Pellegrini, M., Jacobsen, S.E., and Reik, W. (2010). Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. Nature *463*, 1101–1105.

Prakash, S., Johnson, R.E., and Prakash, L. (2005). Eukaryotic translesion synthesis DNA polymerases: specificity of structure and function. Annu. Rev. Biochem. *74*, 317–353.

Prasad, R., Beard, W.A., Batra, V.K., Liu, Y., Shock, D.D., and Wilson, S.H. (2011). A review of recent experiments on step-to-step “hand-off” of the DNA intermediates in mammalian base excision repair pathways. Mol Biol *45*, 536–550.

Prasad, R., Shock, D.D., Beard, W.A., and Wilson, S.H. (2010). Substrate channeling in mammalian base excision repair pathways: passing the baton. Journal of Biological Chemistry *285*, 40479–40488.

Prosperi, E. (2006). The fellowship of the rings: distinct pools of proliferating cell nuclear antigen trimer at work. The FASEB Journal *20*, 833–837.

Rai, K., Huggins, I.J., James, S.R., Karpf, A.R., Jones, D.A., and Cairns, B.R. (2008). DNA demethylation in zebrafish involves the coupling of a deaminase, a glycosylase, and gadd45. Cell *135*, 1201–1212.

Räschle, M., Knipscheer, P., Enoiu, M., Angelov, T., Sun, J., Griffith, J.D., Ellenberger, T.E., Schärer, O.D., and Walter, J.C. (2008). Mechanism of replication-coupled DNA interstrand crosslink repair. Cell *134*, 969–980.

Romanowski, P., Madine, M.A., Rowles, A., Blow, J.J., and Laskey, R.A. (1996). The *Xenopus* origin recognition complex is essential for DNA replication and MCM binding to chromatin. Curr. Biol. *6*, 1416–1425.

Rupp, R.A., and Weintraub, H. (1991). Ubiquitous MyoD transcription at the midblastula transition precedes induction-dependent MyoD expression in presumptive mesoderm of *X. laevis*. Cell *65*, 927–937.

Saha, P., Chen, J., Thome, K.C., Lawlis, S.J., Hou, Z.H., Hendricks, M., Parvin, J.D., and Dutta, A. (1998). Human CDC6/Cdc18 associates with Orc1 and cyclin-cdk and is selectively eliminated from the nucleus at the onset of S phase. Molecular and Cellular Biology *18*, 2758–2767.

Saito, Y., Ono, T., Takeda, N., Nohmi, T., Seki, M., Enomoto, T., Noda, T., and Uehara, Y. (2012). Embryonic lethality in mice lacking mismatch-specific thymine DNA glycosylase is partially prevented by DOPS, a precursor of noradrenaline. Tohoku J. Exp. Med. *226*, 75–83.

Sale, J.E., Lehmann, A.R., and Woodgate, R. (2012). Y-family DNA polymerases and their role in tolerance of cellular DNA damage. Nature Publishing Group *13*, 141–152.

Salguero, I., Guarino, E., Shepherd, M.E.A., Deegan, T.D., Havens, C.G., MacNeill, S.A., Walter, J.C., and Kearsey, S.E. (2012). Ribonucleotide reductase activity is coupled to DNA synthesis via proliferating cell nuclear antigen. Curr. Biol. *22*, 720–726.

Sambrook, J., and Russell, D.W. (2001). Molecular cloning: a laboratory manual (Cold Spring Harbor Press).

Sclafani, R.A., and Holzen, T.M. (2007). Cell cycle regulation of DNA replication. Annu. Rev. Genet. *41*, 237–280.

Shibutani, S.T., la Cruz, de, A.F.A., Tran, V., Turbyfill, W.J., III, Reis, T., Edgar, B.A., and Duronio, R.J. (2008). Intrinsic negative cell cycle regulation provided by PIP box- and Cul4$^{Cdt2}$-mediated destruction of E2F1 during S phase. Developmental Cell *15*, 890–900.

Simonsson, S., and Gurdon, J. (2004). DNA demethylation is necessary for the epigenetic reprogramming of somatic cell nuclei. Nat Cell Biol *6*, 984–990.

Singh, S.P., Lipman, J., Goldman, H., Ellis, F.H., Aizenman, L., Cangi, M.G., Signoretti,

S., Chiaur, D.S., Pagano, M., and Loda, M. (1998). Loss or altered subcellular localization of p27 in Barrett's associated adenocarcinoma. Cancer Research *58*, 1730–1735.

Sive, H.L., Grainger, R.M., and Harland, R.M. Early development of *Xenopus laevis*: A laboratory manual (Cold Spring Harbor Press).

Sjolund, A.B., Senejani, A.G., and Sweasy, J.B. (2012). MBD4 and TDG: Multifaceted DNA glycosylases with ever expanding biological roles. Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis 1–14.

Smet-Nocca, C., Wieruszeski, J.-M., Léger, H., Eilebrecht, S., and Benecke, A. (2011). SUMO-1 regulates the conformational dynamics of thymine-DNA Glycosylase regulatory domain and competes with its DNA binding activity. BMC Biochemistry *12*, 4.

Solnica-Krezel, L., and Sepich, D.S. (2012). Gastrulation: making and shaping germ layers. Annu. Rev. Cell Dev. Biol. *28*, 687–717.

Soucy, T.A., Smith, P.G., and Rolfe, M. (2009). Targeting NEDD8-activated Cullin-RING ligases for the treatment of cancer. Clinical Cancer Research *15*, 3912–3916.

Stokes, M.P. (2002). DNA replication is required for the checkpoint response to damaged DNA in *Xenopus* egg extracts. The Journal of Cell Biology *158*, 863–872.

Stokes, M.P. (2003). DNA damage-induced replication arrest in *Xenopus* egg extracts. The Journal of Cell Biology *163*, 245–255.

Sutton, M.D. (2010). Coordinating DNA polymerase traffic during high and low fidelity synthesis. BBA - Proteins and Proteomics *1804*, 1167–1179.

Suzuki, M.M., and Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet *9*, 465–476.

Tada, S., Li, A., Maiorano, D., Méchali, M., and Blow, J.J. (2001). Repression of origin assembly in metaphase depends on inhibition of RLF-B/Cdt1 by geminin. Nat Cell Biol *3*, 107–113.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., et al. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science *324*, 930–935.

Tardat, M., Brustel, J., Kirsh, O., Lefevbre, C., Callanan, M., Sardet, C., and Julien, E. (2010). The histone H4 Lys 20 methyltransferase PR-Set7 regulates replication origins in mammalian cells. Nat Cell Biol *12*, 1086–1093.

Tutter, A.V., and Walter, J.C. (2006). Chromosomal DNA replication in a soluble cell-free system derived from *Xenopus* eggs. Methods Mol. Biol. *322*, 121–137.

Veenstra, G.J., Destrée, O.H., and Wolffe, A.P. (1999). Translation of maternal TATA-binding protein mRNA potentiates basal but not activated transcription in *Xenopus* embryos at the midblastula transition. Molecular and Cellular Biology *19*, 7972–7982.

Walter, J., and Newport, J. (2000). Initiation of eukaryotic DNA replication: origin unwinding and sequential chromatin association of Cdc45, RPA, and DNA polymerase alpha. Molecular Cell *5*, 617–627.

Walter, J., Sun, L., and Newport, J. (1998). Regulated chromosomal DNA replication in the absence of a nucleus. Molecular Cell *1*, 519–529.

Walter, J.C., and Newport, J.W. (1997). Regulation of replicon size in *Xenopus* egg extracts. Science *275*, 993–995.

Watase, G., Takisawa, H., and Kanemaki, M.T. (2012). Mcm10 Plays a Role in Functioning of the Eukaryotic Replicative DNA Helicase, Cdc45-Mcm-GINS. Curr. Biol. *22*, 343–349.

Waters, T.R., Gallinari, P., Jiricny, J., and Swann, P.F. (1999). Human thymine DNA glycosylase binds to apurinic sites in DNA but is displaced by human apurinic endonuclease 1. J. Biol. Chem. *274*, 67–74.

Williams, K., Christensen, J., and Helin, K. (2012). DNA methylation: TET proteins—guardians of CpG islands? EMBO Rep *13*, 28–35.

Wilson, S.H., and Kunkel, T.A. (2000). Passing the baton in base excision repair. Nat Struct Biol *7*, 176–178.

Winklbauer, R., Nagel, M., Selchow, A., and Wacker, S. (1996). Mesoderm migration in the *Xenopus* gastrula. Int. J. Dev. Biol. *40*, 305–311.

Wold, M.S. (1997). Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. Annu. Rev. Biochem. *66*, 61–92.

Wormington, W.M., and Brown, D.D. (1983). Onset of 5 S RNA gene regulation during *Xenopus* embryogenesis. Dev. Biol. *99*, 248–257.

Wossidlo, M., Nakamura, T., Lepikhov, K., Marques, C.J., Zakhartchenko, V., Boiani, M., Arand, J., Nakano, T., Reik, W., and Walter, J.O.R. (2011). 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. Nature Communications *2*, 241–248.

Wu, J.Q., and Kornbluth, S. (2008). Across the meiotic divide - CSF activity in the post-Emi2/XErp1 era. Journal of Cell Science *121*, 3509–3514.

Wu, S.C., and Zhang, Y. (2010). Active DNA demethylation: many roads lead to Rome. Nature Publishing Group *11*, 607–620.

Yagi, Y., Ogawara, D., Iwai, S., Hanaoka, F., Akiyama, M., and Maki, H. (2005). DNA polymerases η and κ are responsible for error-free translesion DNA synthesis activity over a cis–syn thymine dimer in *Xenopus laevis* oocyte extracts. DNA Repair *4*, 1252–1269.

Yan, S., and Michael, W.M. (2009). TopBP1 and DNA polymerase alpha-mediated recruitment of the 9-1-1 complex to stalled replication forks: implications for a replication

restart-based mechanism for ATR checkpoint activation. Cell Cycle *8*, 2877–2884.

Yanai, I., Peshkin, L., Jorgensen, P., and Kirschner, M.W. (2011). Mapping gene expression in two *Xenopus* species: evolutionary constraints and developmental flexibility. Developmental Cell *20*, 483–496.

Zhang, L., Lu, X., Lu, J., Liang, H., Dai, Q., Xu, G.-L., Luo, C., Jiang, H., and He, C. (2012). Thymine DNA glycosylase specifically recognizes 5-carboxylcytosine-modified DNA. Nature Publishing Group *8*, 328–330.

Zhang, P., Sun, Y., Hsu, H., Zhang, L., Zhang, Y., and Lee, M.Y. (1998). The interdomain connector loop of human PCNA is involved in a direct interaction with human polymerase delta. J. Biol. Chem. *273*, 713–719.

Zhuang, Z., and Ai, Y. (2010). Processivity factor of DNA polymerase and its expanding role in normal and translesion DNA synthesis. BBA - Proteins and Proteomics *1804*, 1081–1093.